OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: October 23, 2003, 16:43:39 ; Search time 2168 Seconds
(without alignments)
9453.747 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctccagcctctcaccga..........cagaatgcggcgatgatgaccat 501
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: en_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pl:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	135.8	27.1	302650	1	AP005958	AP005958 Bradyrhiz
c	2	97.2	19.4	10902	1	AE011858	AE011858 Xanthomon
	3	92.4	18.4	3526	1	AF273216	AF273216 Rhizobium
c	4	88.8	17.7	28804	1	SSU51197	U51197 Sphingomona
c	5	88.8	17.7	28804	6	AR068625	AR068625 Sequence
	6	87.6	17.5	302022	1	AB012558	AB012558 Xylella f
	7	86.8	17.3	15552	1	AE004049	AE004049 Xylella f
	8	86	17.2	1428	1	NME391263	NME391263 Neisseria
	9	86	17.2	4254	1	AF121772	AF121772 Neisseria
	10	86	17.2	7824	1	NME391260	NME391260 Neisseria
c	11	86	17.2	11381	1	AE002524	AE002524 Neisseria
c	12	86	17.2	326301	1	NMA6Z22491	AL162757 Neisseria
c	13	86	17.2	349980	6	AX044033	AX044033 Sequence
c	14	86	17.2	349980	6	AX044034	AX044034 Sequence
c	15	80.6	16.1	301708	1	AB016792	AB016792 Pseudomon
c	16	80.4	16.0	189050	1	AL646077	AL646077 Ralstonia
	17	69.8	13.9	11180	1	AE013811	AE013811 Yersinia
	18	69.8	13.9	334050	1	AJ414151	AJ414151 Yersinia
c	19	67.4	13.5	296600	1	AP005959	AP005959 Bradyrhiz
	20	62.6	12.5	6645	1	BPCYADE	X14199 B.pertussis
	21	56.8	11.3	7493	1	AF193064	AF193064 Caulobact
	22	56.8	11.3	13759	1	AB005779	AB005779 Caulobact
	23	56.8	11.3	19883	1	AF062345	AF062345 Caulobact
c	24	54.2	10.8	303550	1	SCO939118	AL939118 Streptomy
	25	53.2	10.6	208050	1	AL646083	AL646083 Ralstonia
	26	52.6	10.5	4050	1	AY158231	AY158231 Caulobact
c	27	52	10.4	2000	6	AX655393	AX655393 Sequence
	28	51.6	10.3	5120	1	APCLYI	X61112 A.pleuropne
	29	51.6	10.3	8292	1	APHLYIAD	X68595 A.pleuroppe
	30	51.6	10.3	8370	6	AR091872	AR091872 Sequence
	31	51.6	10.3	8370	1	AR268551	AR268551 Sequence
	32	51.2	10.2	300425	1	AP005038	AP005038 Streptomy
	33	48.6	9.7	348411	1	AP003007	AP003007 Mesorhizo
	34	48.4	9.7	4247	1	ACNAPPAB	M65808 Actinobacil
	35	48	9.6	7413	1	PAAPRAPG	X64558 P.aeruginos
	36	48	9.6	11588	1	AE004554	AE004554 Pseudomon
	37	47.8	9.5	6400	6	E30060	E30060 ABC Transpo
	38	47.8	9.5	11273	1	AB023289	AB023289 Pseudomon
c	39	47.6	9.5	11383	1	AB004761	AB004761 Pseudomon
	40	47.2	9.4	299350	1	SME591786	AL591786 Sinorhizo
	41	47	9.4	18857	1	AE006947	AE006947 Mycobacte
	42	47	9.4	35420	1	MTCY22G10	Z84724 Mycobacteri
c	43	47	9.4	42741	1	MSGY423	AD000014 Mycobacte
	44	47	9.4	324050	1	BX248335	BX248335 Mycobacte
	45	46.8	9.3	76093	1	AF416330	AF416330 Ruegeria

ALIGNMENTS

RESULT 1
AP005958/c
LOCUS   AP005958   302650 bp   DNA   linear   BCT 28-MAR-2003
DEFINITION Bradyrhizobium japonicum USDA 110 DNA, complete genome, section
            24/31.
ACCESSION   AP005958 BA000040
VERSION   AP005958.1 GI:27354550
KEYWORDS   .
SOURCE   Bradyrhizobium japonicum USDA 110
ORGANISM   Bradyrhizobium japonicum USDA 110
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1
AUTHORS   Kaneko,T., Nakamura,Y., Sato,S., Minamisawa,K., Uchiumi,T.,
            Sasamoto,S., Watanabe,A., Idesawa,K., Iriguchi,M., Kawashima,K.,

Pred. No. is the number of results predicted by chance to have a

```
           Kohara,M., Matsumoto,M., Shimpo,S., Tsuruoka,H., Wada,T., Yamada,M.
             and Tabata,S.
  TITLE     Complete genomic sequence of nitrogen-fixing symbiotic bacterium
             Bradyrhizobium japonicum USDA110
  JOURNAL   DNA Res. 9 (6), 189-197 (2002)
  MEDLINE   22484998
  PUBMED    12597275
REFERENCE   2
  AUTHORS   Kaneko,T., Nakamura,Y., Sato,S., Minamisawa,K., Uchiumi,T.,
             Sasamoto,S., Watanabe,A., Idesawa,K., Iriguchi,M., Kawashima,K.,
             Kohara,M., Matsumoto,M., Shimpo,S., Tsuruoka,H., Wada,T., Yamada,M.
             and Tabata,S.
  TITLE     Complete genomic sequence of nitrogen-fixing symbiotic bacterium
             Bradyrhizobium japonicum USDA110 (supplement)
  JOURNAL   DNA Res. 9 (6), 225-256 (2002)
  MEDLINE   22485002
  PUBMED    12597279
REFERENCE   3  (bases 1 to 302650)
  AUTHORS   Kaneko,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research
             Institute, The First Laboratory for Plant Gene Research; 2-6-7
             Kazusa-kamatari, Kisarazu, Chiba 292-0812, Japan
             (E-mail:kaneko@kazusa.or.jp),
             URL:http://www.kazusa.or.jp/rhizobase/,
             Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1..302650
                     /organism="Bradyrhizobium japonicum USDA 110"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:224911"
     gene            31..1110
                     /gene="blr6272"
     CDS             31..1110
                     /gene="blr6272"
                     /note="AraC family"
                     ORF_ID:blr6272"
                     /codon_start=1
                     /transl_table=11
                     /product="transcriptional regulatory protein"
                     /protein_id="BAC51537.1"
                     /db_xref="GI:27354551"
                     /translation="MENLLESLVESFGLGGLPVAASGRDAKRGRAMFTDLQAAPAWID
                     LQAEPRDHQLRQAAQAIPRERGWRRIDRGPEEAAIEDITISRWEDFRDSSHDAATTPED
                     RYFVGIALKATRAKLTRDRQVIFDGTMPAGTLYVSAPSKPLAAQFQAPCAFLHFHISA
                     DHFPAQSPAAEGLGDLVLRLPLAAELARALTEHGDAADREFTRCVGQTLAMHLARLE
                     PPRARVNALPKWRLRRVEQYIADNFDRCISLSELANVAGLSRMHFAAQFRATGYRPR
                     EYLLNHRIEHAKTLLATTGRPLAEIALAVGFSTQAHFSTVFKRISGQSPARWRLASKA
                     EPLAAEALPRRRPAADRDWMASAAA"
     gene            complement(1428..3794)
                     /gene="bll6273"
     CDS             complement(1428..3794)
                     /gene="bll6273"
                     /note="ORF_ID:bll6273"
                     /codon_start=1
                     /transl_table=11
                     /product="probable biotin sulfoxide reductase"
                     /protein_id="BAC51538.1"
                     /db_xref="GI:27354552"
                     /translation="MDRYTATMYRNKDQMMDDTIGFPDPGLDLSDGFKPHTSHWGVFS
                     ARNGAAGLEVRAYAGDPDPNGIIDNFPGALRHQARIAQPAIRRGWILERGPGPDDRRGR
                     DEFVSVSWEKALDLLGDELGRIRDTRGPAVFGSYGWSSAGRFHHAQSQVHRFLNIA
                     MGGYVRSVNSYSSGASSVLLPQILAGYEDITKRNVTWEQIATETDIVLAFGGMALKNS
                     MVAGGSISKHVERSAMEAARKRGCEFIIVSPLREDLPVEAGAEWMTCVPSTDTALMLG
                     IVHTLVSEKLHDQAFLDRYTEGSDQPKSAEWAAAISGVDADTIRTL
                     ARRLAGKRALITVSHSLQRAEHGEQPVWMGMVLAAALGQIGLPGGGYAYSLGAIGYYG
                     RRVNDVPGPTLGQGRNGVRDFIPVARIADMLLNPGATYRYNGETRTYPDIRLVYWAGG
                     NPFHHQDINRLRKAFAQLDTFVHELAWTATARHADIVLPSTMTLEREDIGYSSNDP
                     LMVAMHRIAEPFGLARDDYEIFADLAERLGAREPFTEGRTSREWLKHLYEPTRASLAM
                     RGLEAPSPDEFWQRGSLVVPQQPDDGGRLRRFREDPVAHALPTPSGRIEIFSTRIAGH
                     GDADCPGHPVWLDKTDMPKPGAPCFLVANQPVTRLHSQLDFGHSLGAKHRGREVARM
                     NPVDADARGIKDGDIIRLFNERGACLAAVHVTDGIAPGVVQLPTGAWYDPMDPEDDAP
```

```
                     LCVHGNPNVLTRDVGTSSFAQGCTGQLTTVEVBKPTGNLPPIRAFDPV"
     gene            complement(3898..4947)
                     /gene="bll6274"
     CDS             complement(3898..4947)
                     /gene="bll6274"
                     /note="ORF_ID:bll6274"
                     putative dehydrogenase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51539.1"
                     /db_xref="GI:27354553"
                     /translation="MTGASDTVAIGADALRPFAADLLRAGGFTADQADETAAMLVWAN
                     LRGADSHGVLRIPRYVEMVELGLINPAAEPRQVSGKGAVAVIDAQRAPAGAVGMSLAAR
                     KASELADTHGIGWCAVRSISHAGAIGYYAHQVAQGRIGIVMTASKPLMIYHGSSAEG
                     VSTNPIAIAAPTGNPDRPLLLDMSTASVALGKIMAARDAGQPIPVGWAVDGNGQATTD
                     PREAKAVLPMAGPKGSGLSLMIEVLASVLTANPLIAVALVHGGDPGGNGLVAVLDPSA
                     FGDNFTLEVDRLGAAIKSLPPAAGADSVVLPGERGPEEMERRSRAGIPLARGTVSRLL
                     ALASKLNVAVPEGVA"
     gene            complement(4950..5873)
                     /gene="bll6275"
     CDS             complement(4950..5873)
                     /gene="bll6275"
                     /note="ORF_ID:bll6275"
                     hypothetical protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51540.1"
                     /db_xref="GI:27354554"
                     /translation="MSAPPLSQPTREIPMOEKHTFRSDGLKLSAVLHVPDARKPGQRL
                     PAFIVLHGFVGSKDBSHAEIQARMLEQMGIVALRFDFRCCGESEGERAQVRCFDQVAD
                     AKNALTFLABREEVDPGRIGVVGHSFGAAVSVYAAGVDDRFACVISSCGWDGERKFR
                     GQHPTPEAWAKFTSLLEKGRKQKQETGQSMWISRFDAVPIPEHLRKNLSPKAIMEIPV
                     ETAWSMYNFRADEVVGNIARPILFLHTANDTITPTEQSIRMFEKAGQPAELVLTGT
                     SHFPLAEGDAPRTKAIIKGWIDKFPSPLGA"
     gene            complement(5910..6812)
                     /gene="bll6276"
     CDS             complement(5910..6812)
                     /gene="bll6276"
                     /note="ORF_ID:bll6276"
                     unknown protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51541.1"
                     /db_xref="GI:27354555"
                     /translation="MPPIVHAVGHVKINTAIEAVRDATEILGLQITYDEGQTWLS
                     SNGRAAELVLLRSHENATHTIGLEALTVDAVREASRVESAGCRVISRKPSLDCMAAG
                     VTFDTPEGLRPEVHTPTRDQIYGRRYPTHGVGPNRMDHLNLTSPDPIATRQQLERICG
                     LRISRMVDDSLSMWYGGNRQHHIVGVRGKVGLHHYSFEFKBFNDYLKLGDLLDRFD
                     RQMLWGPGRHRPGDNTYAYYTDSSGAMIECSGSMALIADDDGFVPNVITNLERPGNVR
                     CMNVWGTPAPLEWRQFHPFTAIT"
     gene            7080..7853
                     /gene="blr6277"
     CDS             7080..7853
                     /gene="blr6277"
                     /note="GntR family
                     ORF_ID:blr6277"
                     /codon_start=1
                     /transl_table=11
                     /product="transcriptional regulatory protein"
                     /protein_id="BAC51542.1"
                     /db_xref="GI:27354556"
                     /translation="MILTGSRGMAGRRRMRERVKPAGETRAADVLHRMRADIISCTLK
                     PGAKLRFEALRDIYAVSFSTLREALSRLVAEGLVIAEDORGFVVAPVSIDLNDLTYV
                     RVLIRECVALAIKNGDDAWEADIIGAFHRMDRLQNRLGSSNYYLSEEWGKLHGDFHFS
                     LVAACGSPNLLEIRQKLFERAHRYRRMSSQFRTKWRAKDVEHKMIMDSVVARDAAKAE
                     ELIERHIRETTDNVIKHAGHLFVTSDEERPRRTSRVAAE"
     gene            complement(7993..8880)
                     /gene="ilvE"
     CDS             complement(7993..8880)
                     /gene="ilvE"
                     /note="ORF_ID:bll6278"
                     /codon_start=1
                     /transl_table=11
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: October 23, 2003, 17:23:09 ; Search time 1656 Seconds
                                (without alignments)
                                7352.985 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctcagcctccaccga........cagaatgcggcgatgatcat 501

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_estom:*
                 17: em_gss_hum:*
                 18: em_gss_inv:*
                 19: em_gss_pln:*
                 20: em_gss_vrt:*
                 21: em_gss_mam:*
                 22: em_gss_mus:*
                 23: em_gss_pro:*
                 24: em_gss_rod:*
                 25: em_gss_phg:*
                 26: em_gss_vri:*
                 27: gb_gss1:*
                 28: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES

         %
Result   Query   Description
No.   Score Match Length DB ID
---- ----- ----- ------ -- -- ---------------   --------------
c 1 60.6 12.1   932   29 CNS0070E   AL066254 Drosophil
c 2 55   11.0   925   29 CNS0091P   AL053013 Drosophil
c 3 49.4   9.9   812   29 BZ549462   BZ549462 pacs1-60
c 4 45.2   9.0 1148   13 BX374895   BX374895 BX374895

c   5   44.8 8.9 508   9   AV630811   AV630811 AV630811
c   6   43.6 8.7 772 28   AF075872   AF075872 AF075872
c   7   42   8.4 839 29   CNS004NB   AL054280 Drosophil
c   8   41.4 8.3 612 14   CB214812   CB214812 OML05092
c   9   41.4 8.3 620 14   CB211981   CB211981 OML02261
c 10   41   8.2 1201 13   BX360624   BX360624 BX360624
c 11   40.8 8.1 932 29   CNS0072Q   AL066742 Drosophil
c 12   40.8 8.1 1201 13   BX355664   BX355664 BX355664
c 13   40.6 9   426   9   AJ474950   AJ474950 AJ474950
c 14   40.6 8.1 835 13   BX437276   BX437276 BX437276
c 15   40.6 8.1 851 29   BZ557064   BZ557064 pacs1-60
c 16   40.6 8.1 919 29   BZ551811   BZ551811 pacs1-6C
c 17   40.6 8.1 927 29   BQ890839   BQ890839 AGENCOURT
c 18   40.6 8.1 774 29   BZ579186   BZ579186 msh2_6237
c 19   40.4 8.1 478 13   BI956591   BI956591 HVSMEn000
c 20   40.2 8.0 504 12   BU980310   BU980310 HA20E19r
c 21   40.2 8.0 531 14   BJ450136   BJ450136 BJ450136
c 22   40.2 8.0 581 9   CD054442   CD054442 HO01G23r
c 23   40.2 8.0 586 9   AV921113   AV921113 AV921113
c 24   40.2 8.0 588 9   AV913605   AV913605 AV913605
c 25   40.2 8.0 605 9   BJ468187   BJ468187 BJ468187
c 26   40.2 8.0 626 12   AV918908   AV918908 AV918908
c 27   40.2 8.0 628 12   BJ473505   BJ473505 BJ473505
c 28   40.2 8.0 638 14   BJ457563   BJ457563 BJ457563
c 29   40.2 8.0 659 9   CB880020   CB880020 HM05D24w
c 30   40.2 8.0 660 9   AV931532   AV931532 AV931532
c 31   40.2 8.0 660 12   AV946608   AV946608 AV946608
c 32   40.2 8.0 660 12   BJ465650   BJ465650 BJ465650
c 33   40.2 8.0 588 12   BJ475855   BJ475855 BJ475855
c 34   39.8 7.9 603 13   BU035564   BU035564 946128A12
c 35   39.8 7.9 627 13   BQ279904   BQ279904 1091033D1
c 36   39.8 7.9 633 13   BU092391   BU092391 946130C06
c 37   39.8 7.9 639 13   BU098243   BU098243 946133A10
c 38   39.6 7.9 574 12   AI770858   AI770858 606059D06
c 39   39.6 7.9 576 9   BM349363   BM349363 MEST248-F
c 40   39.6 7.9 577 12   AI881920   AI881920 606075B06
c 41   39.6 7.9 643 13   BM267007   BM267007 MEST389-A
c 42   39.6 7.9 645 29   BU499802   BU499802 946179C03
c 43   39.6 7.9 673 29   CNS012I3   AL101589 Drosophil
c 44   39.6 7.9 686 12   BZ563832   BZ563832 pacs2-164
c 45   39.6 7.9   BI959511   BI959511 HVSMEn001

                           ALIGNMENTS

RESULT 1
CNS0070E/c
LOCUS   CNS0070E   932 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TBT3 end of BAC #
            BAC14D21 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL066254
VERSION   AL066254.1 GI:4945121
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 932)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Oseogawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1..932
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR14D21"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT       421 a    135 c    96 g    50 t    230 others
ORIGIN

Query Match      12.1%;  Score 60.6; DB 29; Pred. No. 3.8e-05; Length 932;
Best Local Similarity   29.7%;  Indels 0; Gaps 0;
Matches 77;  Conservative 83;  Mismatches 99;

Qy     47 GCGGTCGGGGGCAAAACCCGGACGAGCTGGCTTCCTGCCCGGCCGCCTCCGAAATTGTC 106
Db    890 GBTSTYGGSGSSCSYGCKCKGBSSTGYCKCGBCYSYSSBKCKBYGGBYKCKCGCBGYSS 831
Qy    107 GAGACGCGCCCATCTCCACCGCGAGACTCACGGCCGCCGCCTTGCTTGCTGCCTTGTTCTAC 166
Db    830 GCKSBGCCCSGCBCBGCYSCBSSGCGTSSTCYTYKCCGGYSBCGKCTGYKCKCTYKYKKGYK 771
Qy    167 TGGCCCGTGGCGTGGCGTGKSGTGKSGTTBGCGGGTKCKYTBTTTGKTBGTGYTKKYSGCKCSYGKBTCTYGCCBS 226
Db    770 GGTGKSGTGKSGTTBGCGGGTKCKYTBTTTGKTBGTGYTKKYSGCKCSYGKBTCTYGCCBS 711
Qy    227 ATCGTGCCGGGCGACCGTGTAAAGCTGGTTCAGCCGTGCGGCTCGGAGGTCGGTGTGCGGGGC 286
Db    710 YGYSCKCTGCKCGSCKCSTCKCKCTCBCGSTSSYGCTSSYGCTCGYTGYSTYTC 651
Qy    287 ACTCATGTCCCGATGGCC 305
Db    650 TYTKTTYYCCTSYCBCGCY 632

RESULT 2
CNS0091P
LOCUS       CNS0091P      925 bp    DNA    linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1   GI:4934461
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 925)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
```

EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1..925
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR19D16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT       120 a    61 c    61 g    172 t    511 others
ORIGIN

Query Match      11.0%;  Score 55; DB 29; Pred. No. 0.001; Length 925;
Best Local Similarity   14.8%;  Mismatches 145;  Indels 6;  Gaps 1;
Matches 57;  Conservative 177;

Qy     31 GGCTCCGTGGAAGTGGTTGCGTCGGCGGGGCAAAACCCGGACGAGCTGGCCTTCCTGCCCGGC 90
Db    539 KKCGSTTBGSTTTTTTSSGSGYGKGCSSGSGBSCSCCSCSSCSSSCSCCSBCCCCCSCSS 598
Qy     91 CGCCCTCGAAATTGTCGAGACGCGCCCATCTCCACCGCGAGACTCACGGCCGCCGCCTTGCT 150
Db    599 YCCSSBSSSKCCSSTSBSCSCCCSSKSVCGTSCSSSSSCSCSSSSSTSSSSTSSSSTSSKSSS 658
Qy    151 TGCTGCCTTGTTCTACTGCCGCCTGCGGTCTCGGCGGGTCTCGGCAGGATCGACATCGTTGC 210
Db    659 SGSSSSSSSYYTTSKSTSASGSGSWSAGGGSGSTGSTSSSSSSSTS*SSSVSSGSKSS 718
Qy    211 TTCTGCATCAGAAAGATC------GTGCCGGGCGACCGTGTAAAGCTGGTTCAGCCGCT 264
Db    719 TBSSGSBBSSGSSGSSSSTSSSSSSSSBBCSTSSSSSSSYSSTCSCCSTCCSYSYSSSTSSSSS 778
Qy    265 CGAGGTCGGCGTGTGCGGGGCACTCATGTCCGCGATGGCCAAACCGTCAAGGCCGGCGA 324
Db    779 TSWGSTSGSSSSVGTSSSSSVTSSSSDSTSTCCSCCCYMCTCCSTYBMBCYTSTSCGGSSSSSGKG 838
Qy    325 GATTCTGATCGAGCTGGATCCATTCGGGGTGGTGTGCGGGTGTGTGATGTYGCGCCCGTCAGAGGTC 384
Db    839 GVTKCGCGCGCGSSSTNGMBGTSSACSSSGSSSSSSCSSSSVSSSSKSASSSSVSSSGSSSGVSS 898
Qy    385 CATCACGGTGTCGGCGCCCCACCGA 409
Db    899 NSSSASKSSSSSGSVSSGSSGSGGSV 923

RESULT 3
BZ549462
LOCUS       BZ549462      812 bp    DNA    linear   GSS 17-DEC-2002
DEFINITION  pacs1-60_2079.x1 pacs1-60 Pseudomonas aeruginosa genomic clone
            pacs1-60_2079, genomic survey sequence.
ACCESSION   BZ549462
VERSION     BZ549462.1   GI:27153043
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 812)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        October 23, 2003, 16:42:52 ; Search time 232 Seconds
                                 (without alignments)
                                 5829.390 Million cell updates/sec

Title:         US-09-484-577A-3
Perfect score: 501
Sequence:      1 actcccagcctctcaccga.........cagaatgcggcgatgatcat 501

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      252756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters:        5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     N Geneseq_19jun03:*
     1: /SIDS17/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
     2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
     3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
     4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
     5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
     6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
     7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
     8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
     9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
    10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
    11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
    12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
    13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
    14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
    15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
    16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
    17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
    18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
    19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
    20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
    21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
    22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
    23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
    24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
    25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

                        SUMMARIES
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.		Score	% Match	Query Length	DB ID		Description
c	1	88.8	17.7	28804	17	AAT37329	Sphingan biosynthe
c	2	88.8	17.7	28804	18	AAT92474	Sphingomonas genus
c	3	88.8	17.7	28804	20	AAV99812	Sphingomonas S88 s
c	4	88.8	17.2	28804	20	AAV81474	Chromosomal fragme
	5	86	17.2	1305	21	AAZ54336	Neisseria meningit
	6	86	17.2	1428	21	AAZ54335	Neisseria meningit
	7	86	17.2	24158	21	AAB81532	N. meningitidis pa
c	8	86	17.2	349980	21	AAF21611	Neisseria meningit

```
c   9   86    17.2  349980  21  AAF21612   Neisseria meningit
c  10   86    17.2  837096  21  AAA81489   N. meningitidis pa
   11   51.6  10.3    5120  18  AAT73218   ApxIBD gene. Acti
   12   51.6  10.3    5120  20  AAA28585   A. pleuropneumonia
   13   47.8   9.5    6400  20  AAA222701  Pseudomonas fluore
   14   47     9.4 4403765  22  AAI99683   Mycobacterium tube
   15   47     9.4 4411529  22  AAI99682   Mycobacterium tube
   16   45     9.0    7184  15  AAQ70050   ptx gene of Pasteu
   17   44.8   8.9  349980  21  ABQ81842   Bifidobacterium lo
   18   43.4   8.7    7183  11  AAQ06074   Sequence encoding
   19   43     8.6   24379  18  AAT93095   Streptomyces freno
   20   43     8.6   24379  19  AAV25925   Streptomyces roseo
c  21   42     8.4 4403765  22  AAI99683   Mycobacterium tube
c  22   42     8.4 4411529  22  AAI99682   Mycobacterium tube
c  23   40     8.0   45055  25  AB266808   Orthosomycin biosy
c  24   39.8   7.9    1320  25  ABK56068   M. echinospora cal
c  25   39     7.8    1761  22  AAH66404   C glutamicum codin
   26   39     7.8    1998  22  AAF72013   Corynebacterium gl
   27   39     7.8  349980  22  AAH68528   C glutamicum codin
   28   38.2   7.6    5802  24  ABS78696   S. kaniharaensis D
   29   37.6   7.5    1553  23  AAS89358   DNA encoding novel
c  30   37.6   7.5    2931  23  AAS92554   DNA encoding novel
c  31   37.6   7.5   10732  21  AAA10594   Gene encoding a su
c  32   37.2   7.4     375  21  ABX48203   Bovine EST associa
   33   37.2   7.4    1194  17  AAT58555   Streptomyces prist
   34   37.2   7.4    4496  17  AAT58553   Streptomyces prist
   35   37.2   7.4   58857  17  AAA58471   Nucleotide sequenc
c  36   37.2   7.4   76804  24  ABS78942   E. coli CFT073 gen
   37   37.2   7.4  349980  24  ABQ81847   Bifidobacterium lo
c  38   37     7.4    1951  21  AAV11459   C. acidivorans gam
c  39   37     7.4    1981  21  AAZ50482   Corn sulphate perm
   40   36.6   7.3     276  24  ABN75424   Human transport pr
   41   36.4   7.3    1408  24  ABK63812   Rat sequence diffe
   42   36.4   7.3    1739  22  AAD19384   Rat CCAAT/enhancer
   43   36.4   7.3    1739  23  ABK51406   DNA encoding rat C
   44   36.4   7.3    1739  23  AAL44082   DNA encoding rat C
   45   36.4   7.3    1739  24  ABK51998   DNA encoding wild-
```

                                 ALIGNMENTS

```
RESULT 1
AAT37329/c
ID   AAT37329 standard; DNA; 28804 BP.
XX
AC   AAT37329;
XX
DT   30-NOV-1996  (first entry)
XX
DE   Sphingan biosynthetic gene region.
XX
KW   Sphingan; polysaccharide; spsB gene; glucosyl-1P-transferase; ds.
XX
OS   Sphinomonas strain S88 (ATCC 31554).
XX
FH   Key             Location/Qualifiers
FT   CDS             complement (1942..1944)
FT                   /*tag= a
FT                   /codon_start= 1942..1944
FT                   /note= "spsG gene putative initiation codon"
FT   CDS             complement (3311..3313)
FT                   /*tag= b
FT                   /codon_start= 3311..3313
FT                   /note= "spsS gene putative initiation codon"
FT   CDS             complement (5323..5325)
FT                   /*tag= c
FT                   /codon_start= 5323..5325
FT                   /note= "spsR gene putative initiation codon"
FT   CDS             5526..5528
FT                   /*tag= d
FT                   /codon_start= 5526..5528
FT                   /note= "spsQ gene putative initiation codon"
```

CDS

FT complement (7076..7078)
FT /*tag= e
FT /codon_start= 7076..7078
FT /note= "spsI gene putative initiation codon"
FT 7588..7590
FT /*tag= f
FT /codon_start= 7588..7590
FT /note= "spsK gene putative initiation codon"
FT 8643..8645
FT /*tag= g
FT /codon_start= 8643..8645
FT /note= "spsL gene putative initiation codon"
FT complement (10938..10940)
FT /*tag= h
FT /codon_start= 10938..10940
FT /note= "spsJ gene putative initiation codon"
FT 11569..11571
FT /*tag= i
FT /codon_start= 11569..11571
FT /note= "spsF gene putative initiation codon"
FT 12886..12888
FT /*tag= j
FT /codon_start= 12886..12888
FT /note= "spsD gene putative initiation codon"
FT 15165..15167
FT /*tag= k
FT /codon_start= 15165..15167
FT /note= "spsE gene putative initiation codon"
FT 15883..15885
FT /*tag= l
FT /codon_start= 15883..15885
FT /note= "urf32 gene putative initiation codon"
FT 16748..16750
FT /*tag= m
FT /codon_start= 16748..16750
FT /note= "urf26 gene putative initiation codon"
FT complement (21082..21084)
FT /*tag= n
FT /codon_start= 21082..21084
FT /note= "atrB gene putative initiation codon"
FT 21706..23118
FT /*tag= o
FT /label= spsB
FT /product= glucosyl IP-transferase
FT 23238..23240
FT /*tag= p
FT /codon_start= 23238..23240
FT /note= "rhsA gene putative initiation codon"
FT 24113..24115
FT /*tag= q
FT /codon_start= 24113..24115
FT /note= "rhsC gene putative initiation codon"
FT 24683..24685
FT /*tag= r
FT /codon_start= 24683..24685
FT /note= "rhsB gene putative initiation codon"
FT 25744..25746
FT /*tag= s
FT /codon_start= 25744..25746
FT /note= "rhsD gene putative initiation codon"
FT complement (27534..27536)
FT /*tag= t
FT /codon_start= 27534..27536
FT /note= "urf31 gene putative initiation codon"
FT 27747..27749
FT /*tag= u
FT /codon_start= 27747..27749
FT /note= "urf34 gene putative initiation codon"
XX
PN EP728841-A2.
PD 28-AUG-1996.
XX

PF 24-JAN-1996; 96EP-0300467.
XX
PR 24-JAN-1995; 95US-0377440.
XX
PA (SHIN-) SHINETSU BIO INC.
PA (SHIE) SHINETSU CHEM CO LTD.
XX
PI Armentrout RW, Mikolajczak M, Pollock TJ, Thorne L;
PI Yamazaki M;
XX
DR WPI: 1996-386292/39.
DR P-PSDB; AAW03997.
XX
PT New isolated DNA from Sphingomonas sp. - used for transforming
PT recipient bacteria to obtain hyper-producers of sphingan
PT polysaccharide(s).
XX
PS Claim 32; Page 56-70; 105pp; English.
XX
CC A 28.8 kb chromosomal fragment of Sphingomonas strain S88 was
CC isolated on the basis of its ability to restore sphingan
CC biosynthetic capability to Sphingomonas mutant S88m260. It
CC contains 23-25 genes, including sps genes coding for biosynthesis of
CC the polysaccharide sphingan, rhs genes coding for dTDP-(L)rhamose
CC biosynthesis, atrDB genes coding for a transport function and some
CC unidentified open translation reading frames (urf). The spsB gene
CC was identified that is believed to code for glucosyl IP-transferase
CC (AAW03997), an enzyme catalysing the first step of assembly of
CC sphingan carbohydrates. DNA fragments of S88 can be inserted into
CC a vector in multiple copies and used to produce engineered bacteria
CC that are hyper-producers of sphingan.
XX
SQ Sequence 28804 BP; 4974 A; 9806 C; 9228 G; 4796 T; 0 other;

Query Match   17.7%;   Score 88.8;   DB 17;   Length 28804;
Best Local Similarity   56.5%;   Pred. No. 9.4e-13;
Matches 165;   Conservative   0;   Mismatches 127;   Indels   0;   Gaps   0;

Qy   80 TTCCTGCCGGCCGGCCGCCTCCGAAATTGTCGAGACGCGCCATCTCCCACCGGAGACTCACG 139
          ||||||||| |||| |||| ||||| ||||||||| || |||||||||||| ||||||
Db 18808 TTCCTGCCGGCCGGCCGCCTCGAGATCATCGAACGTCCGGTGTCGCCCACCGGCGCCCTCACC 18749

Qy   140 GCCGCCTTGCTGCCCTGTTCTACTGCGCGGGGTGGCGGGGTCTCGGCAGGATC 199
          || |||| |||| ||| || |||||| ||||
Db 18748 GGCCGGGTGATCGGTGGCGGGGCTGGCGGATCAATCGGCCCGCGTG 18689

Qy   200 GACATCGTTGCTTCTGCATCCAGAAAGATCGTGCCGGGCGACCGTGTAAAGCTGGTTCAG 259
          || ||||| ||| |||| ||||||| || |||| ||||||
Db 18688 GAAGTCGTGGCGGGCCCGACCATCGGCCTGGCTGCGAGCAAGATCGTGCGCAG 18629

Qy   260 CCGCTCGAGGTCGGCGGCGTGGTGCGGGGCCACTCATTCGGGGCGGATGGCCAAACCGTCAAGGCC 319
          ||||| ||||| |||| |||||| || ||| |||| ||||
Db 18628 TCCCCCGAAAGCGGTATCGTCCGGCGCATTCTGGTTGGGCCAGGGCGCAGGGGCAGAAGGTTCAGAAG 18569

Qy   320 GGCGGAGATTCTGATCGAGCTCGATCCATTCCGCGGGGTGGTGTGGATGTTGCGC 371
          ||| ||||| |||||| ||||||||| |||| ||||| |||| |||||| |||
Db 18568 GGCCAGGTGCTGATCACGCTCGATCCGCATCCCACCATGTCGCAGCCGAAGCCGCGC 18517

RESULT 2
AAT92474/c
ID AAT92474 standard; DNA; 28804 BP.
XX
AC AAT92474;
XX
DT 04-FEB-1998 (first entry)
XX
DE Sphirgomonas genus microbe isolated DNA sequence producing sphingan.
XX
KW Sphingomonas microbe; sphingan polysaccharide biosynthesis gene;
KW sphingan S-88; spsB gene; ss.
XX
OS Sphingomonas sp.

```
OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003 Compugen Ltd.

Run on:      October 23, 2003, 18:10:19  ; Search time 222 Seconds
                                         (without alignments)
                                         6052.172 Million cell updates/sec

Title:          US-09-484-577A-3
Perfect score:  501
Sequence:       1 acctccagcctccaccga........cagaatgcggcgatgatcat 501

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters:      3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published Applications NA:*
          1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
          2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
          3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
          4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
          5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
          6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
          7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
          8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
          9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
         10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
         11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
         12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
         13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
         14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
         15:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
         16:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
         17:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Match	Query Length	DB	ID	Description	
1		59.2	11.8	1428	11	US-09-884-696-36	Sequence 36, Appl	
2		51.2	10.2	1803	14	US-10-156-761-4249	Sequence 4249, Ap	
3		51.2	10.2	9025608	14	US-10-156-761-1	Sequence 1, Appli	
4		43.8	8.7	819	14	US-10-156-761-4175	Sequence 4175, Ap	
5		43.4	8.7	681	14	US-10-156-761-2703	Sequence 2703, Ap	
6	c	43.4	8.7	9025608	14	US-10-156-761-1	Sequence 1, Appli	
7		40.6	8.1	768	14	US-10-156-761-6013	Sequence 6013, Ap	
8	c	40.6	8.1	1758	14	US-10-156-761-6524	Sequence 6524, Ap	
9		39.4	7.9	393	14	US-10-156-761-1979	Sequence 1979, Ap	
10	c	39	7.8	1761	10	US-09-738-626-1439	Sequence 1439, Ap	
11		39	7.8	3309400	10	US-09-738-626-1	Sequence 1, Appli	
12		38.4	7.7	1050	14	US-10-156-761-4467	Sequence 4467, Ap	
13		38.2	7.6	1578	14	US-10-156-761-6129	Sequence 6129, Ap	
14	c	37.8	7.5	1038	14	US-10-156-761-6667	Sequence 6667, Ap	
15	c	37.4	7.5	2715	14	US-10-156-761-2245	Sequence 2245, A	
16	c	37.2	7.4	375	10	US-09-960-352-13368	Sequence 13368, A	
17	c		37.2	1080	14	US-10-156-761-3076	Sequence 3076, Ap	
18			37.2	76804	12	US-10-085-959-109	Sequence 109, App	
19			37	477	14	US-10-156-761-6236	Sequence 6236, Ap	
20			37	2403	14	US-10-156-761-4600	Sequence 4600, Ap	
21			36.8	875	10	US-09-903-814-21	Sequence 21, Appl	
22			36.8	1125	14	US-10-156-761-7230	Sequence 7230, Ap	
23			36.6	1146	14	US-10-156-761-3791	Sequence 3791, Ap	
24	c		36.6	1824	14	US-10-156-761-6480	Sequence 6480, Ap	
25			36.4	1408	14	US-10-156-761-1719	Sequence 1719, Ap	
26	c		36.4	1739	9	US-09-917-800A-1719	Sequence 10, Appl	
27			36.4	1739	12	US-09-789-836-10	Sequence 10, Appl	
28			36.2	1107	14	US-09-789-831-10	Sequence 5378, Ap	
29	c		36	1335	14	US-10-156-761-5378	Sequence 1743, Ap	
30	c		36	88421	10	US-10-156-761-1743	Sequence i, Appli	
31	c		35.8	1371	14	US-09-976-059-1	Sequence 5846, Ap	
32			35.8	3765	12	US-10-156-761-5846	Sequence 9, Appli	
33			35.8	3765	14	US-09-953-348-9	Sequence 9, Appli	
34			35.8	53500	14	US-10-267-255-9	Sequence 76, Appl	
35	c		35.8	53500	14	US-09-953-348-76	Sequence 76, Appl	
36	c		35.8	154746	14	US-10-267-255-76	Sequence 8, Appli	
37	c		35.6	978	12	US-09-827-688-8	Sequence 2643, Ap	
38			35.6	1125	14	US-10-156-761-2643	Sequence 2293, Ap	
39			35.6	45191	15	US-10-156-761-2293	Sequence 649, App	
40	c		35.2	500	12	US-10-080-170-649	Sequence 84, Appl	
41			35.2	1359	14	US-10-063-685-84	Sequence 6824, Ap	
42			35.2	1587	14	US-10-156-761-6824	Sequence 4602, Ap	
43			35.2	2076	14	US-10-156-761-4602	Sequence 3702, Ap	
44			35.2	25165	14	US-10-156-761-3702	Sequence 39, Appl	
45	c		35	843	14	US-10-114-170-39	Sequence 1925, Ap	
						US-10-156-761-1925		

ALIGNMENTS

```
RESULT 1
US-09-884-696-36
; Sequence 36, Application US/09884696
; Publication No. US20031003580A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT: ANGELOS, JOHN A
;  APPLICANT: HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION: BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE: 2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
;  LENGTH: 1428
;  TYPE: DNA
;  ORGANISM: Moraxella bovis
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(1425)
US-09-884-696-36

  Query Match             11.8%;  Score 59.2;  DB 11;  Length 1428;
  Best Local Similarity   49.1%;           Pred. No. 5.2e-08;
  Matches  157;  Conservative    0;  Mismatches 163;  Indels    0;  Gaps     0;

Qy          68 GACGAGCTGGCCTTCCTGCCCGGCCGCCCTCCGAAATTGTCGAGACGCGCCATTCCCACC 127
               ||||||||| ||||| |||| || || |||| ||| | ||| || || ||| |||  |||
Db         103 GAAGAACTCGCTTTCTTCCTGCACATCTGCAAGTCTAGAACTCAGAACACCTGTATCCAGATCT 162

Qy         128 GCGAGACTCACGGCGCGCCTTGCTTCGCTGCCTTGTTCTACTGCGCCGTGGCGCGCGGT 187
               |||| |||| ||| ||   |||  | ||| || || || || |  |||| |||| |||
Db         163 TCTAAGTGGACAGCTAGAATAATCATGATATTTGTCCTATTTGCTTGCTATGGTCTTGG 222

Qy         188 CTCGGCAGGATCGACATCGTTGCTTCTGCATCCAGAAAGATCGTGCCGGGCGACCGTGTA 247
               ||| |   |  |    |  | |  ||  |||| || |||| || || ||| || || |||
```

Db   223 GTTGGACAGATTGACATTGTTGCTACACAGCTTCAGGTAAAATTCTTCAGGTAGCCGTAGC 282

Qy   248 AAGCTCGTTCAGCCGCTCGAGGTCGGCGTGGTGCGGGCCACTCATGTCCGGATGGCCAA   307
           ||| ||||| |||| || | || | |||| |||| || |||| |||
Db   283 AAGACTATTCAATCTTGGAAACAGCGATAGTTAAAGCAGTTTATGTACGTGATGGTCAA   342

Qy   308 ACCGTCAAGGCGCGGCGAGATTCTGATCGGATCGGATCCATTCGCGCGGGTGGTGTGGATGTT   367
           ||| | ||||| || |||| ||| |||||| ||||| |||||
Db   343 AATGTTCAACAAGGTGAAATATTAGTAGATTTAGTGGGAATCGGTTCAGATAGTGATGTT   402

Qy   368 GCGCCCCGTCAGAGGTCCAT   387
           |||| || |||
Db   403 GCTCAGTCCGAGAAAGCCCT   422

RESULT 2
US-10-156-761-4249/c
; Sequence 4249, Application US/10156761
; Publication No. US2003011901801
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4249
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1803)
US-10-156-761-4249:

Query Match   10.2%;   Score 51.2; DB 14; Length 1803;
Best Local Similarity 69.4%;   Pred. No. 4e-05;
Matches 100; Conservative 0; Mismatches 38; Indels 6; Gaps 2;

Qy   333 TCGAGCTGGATCCATTCGCGCGGGTGGTGGATGTTGCGCCCCGTC-AGAGGTCCATCACG   391
           | | ||| || ||||||||||||||||||||||||| ||||| |||| | ||||||||||
Db   764 TTGCGCAGGACCCATTCGCGCGGTGGTGGATGTTGCGGCCGGTGGACAGGTCCATGACG   705

Qy   392 GTGTCGGCGCGCCCACGGATCGCCACCATCTTGTCGACCTTTCTTCACCGACG----A   446
           || ||||||| ||| ||||| || | ||||| ||| |||| ||| || ||
Db   704 GTGTCGGCGCGCCCCAGCGAGGTCGCCCAGGTCATCTTCTCGACCTCCTCCTCCGGATGGAGGAA   645

Qy   447 GTCACCGCCGAGTTGCCGATATTG   470
           ||||| |||| ||||||||||||
Db   644 GTGACCGCGGAGTTGCCGATGTTG   621

RESULT 3
US-10-156-761-1
, Sequence 1, Application US/10156761
, Publication No. US2003011901801
; GENERAL INFORMATION:
, APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1:

Query Match   10.2%;   Score 51.2; DB 14; Length 9025608;
Best Local Similarity 69.4%;   Pred. No. 3.9e-05;
Matches 100; Conservative 0; Mismatches 38; Indels 6; Gaps 2;

Qy   333 TCGAGCTGGATCCATTCGCGCGGGTGGTGGATGTTGCGCCCCGTC-AGAGGTCCATCACG   391
           | | ||| || ||||||||||||||||||||||||| ||||| |||| | ||||||||||
Db 5232924 TTGCGCAGGACCCATTCGCGCGGTGGTGGATGTTGCGGCCGGTGGACAGGTCCATGACG 5232983

Qy   392 GTGTCGGCGCGCCCACGGATCGCCACCATCTTGTCGACCTTTCTTCACCGACG----A   446
           || ||||||| ||| ||||| || | ||||| ||| |||| ||| || ||
Db 5232984 GTGTCGGCGCGCCCCAGCGGGTCGCCCAGGTCATCTTCTCGACCTCCTCCGATGGAGGAA 5233043

Qy   447 GTCACCGCCGAGTTGCCGATATTG   470
           ||||| |||| ||||||||||||
Db 5233044 GTGACCGCGGAGTTGCCGATGTTG 5233067

RESULT 4
US-10-156-761-4175
; Sequence 4175, Application US/10156761
; Publication No. US2003011901801
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4175
; LENGTH: 819
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(819)
US-10-156-761-4175:

Query Match   8.7%;   Score 43.8; DB 14; Length 819;
Best Local Similarity 45.9%;   Pred. No. 0.002;
Matches 150; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy   62 ACCCGGGACCGAGCTGGCCTTCCTGCCGCGCCGCCCTCGAAATTGTCGAGACGCCGCCATCT 121
          ||||||||| ||| | || ||| || | || || | ||| ||| / ||| / / ||

```
OM nucleic - nucleic search, using sw model

                    Copyright   GenCore version 5.1.6
                    Copyright (c) 1993 - 2003  Compugen Ltd.

Run on:             October 23, 2003, 17:26:14  ; Search time 58 Seconds
                                                  (without alignments)
                                                  3812.637 Million cell updates/sec

Title:              US-09-484-577A-3
Perfect score:      501
Sequence:           1 actctccagcctcttcaccga.........cagaatgcggcgatgatcat 501

Scoring table:      IDENTITY NUC
                    Gapop 10.0, Gapext 1.0

Searched:           569978 seqs, 22069l566 residues

Total number of hits satisfying chosen parameters:    :139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
             1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
             2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
             3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
             4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
             5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
             6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                SUMMARIES
```

	Result No.	Score	% Query Match	Match Length	DB	ID	Description
c	1	88.8	17.7	28804	2	US-08-592-874-1	Sequence 1, Appli
c	2	88.8	17.7	28804	2	US-09-096-942-2	Sequence 2, Appli
c	3	88.8	17.7	28804	3	US-09-096-867-2	Sequence 2, Appli
	4	51.6	10.3	5120	3	US-08-772-270A-6	Sequence 6, Appli
	5	51.6	10.3	8370	3	US-08-488-706-1	Sequence 1, Appli
	6	51.6	10.3	8370	4	US-09-062-126-1	Sequence 1, Appli
	7	48	9.6	1380	4	US-09-252-991A-3550	Sequence 3550, Ap
	8	48	9.6	1947	4	US-09-252-991A-3557	Sequence 3557, Ap
c	9	47.6	9.5	852	4	US-09-252-991A-11544	Sequence 11544, A
c	10	47.6	9.5	1608	4	US-09-252-991A-11665	Sequence 11665, A
c	11	47.6	9.5	1608	4	US-09-252-991A-11769	Sequence 11769, A
	12	47	9.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	13	47	9.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	14	42	8.4	4403765	3	US-09-103-840A-1	Sequence 2, Appli
	15	42	8.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	16	40.6	8.1	1572	4	US-09-252-991A-3270	Sequence 3270, Ap
c	17	40.6	8.1	1965	4	US-09-252-991A-3234	Sequence 3234, Ap
c	18	40.6	8.1	2325	4	US-09-252-991A-3413	Sequence 3413, Ap
	19	40.2	8.0	822	4	US-09-252-991A-9029	Sequence 9029, Ap
c	20	40.2	8.0	1527	4	US-09-252-991A-8978	Sequence 8978, Ap
c	21	40.2	8.0	2043	4	US-09-252-991A-8759	Sequence 8759, Ap
c	22	39.2	7.8	2034	4	US-09-252-991A-16314	Sequence 16314, A
c	23	39.2	7.8	2190	4	US-09-252-991A-15964	Sequence 15964, A
c	24	39	7.8	1254	4	US-09-252-991A-4072	Sequence 4072, Ap
	25	39	7.8	1404	4	US-09-252-991A-4118	Sequence 4118, Ap
c	26	39	7.8	1680	4	US-09-252-991A-4087	Sequence 4087, Ap
c	27	38.4	7.7	435	4	US-09-252-991A-81	Sequence 81, Appli
c	28	38.4	7.7	630	4	US-09-252-991A-105	Sequence 105, App
c	29	38.4	7.7	1131	4	US-09-252-991A-76	Sequence 76, Appl
c	30	38.4	7.7	1353	4	US-09-252-991A-88	Sequence 88, Appl
	31	38.4	7.7	1890	4	US-09-252-991A-206	Sequence 206, App
	32	37.2	7.4	1194	4	US-08-765-907A-9	Sequence 9, Appli
c	33	37.2	7.4	4496	4	US-08-765-907A-6	Sequence 6, Appli
c	34	37	7.4	1951	3	US-08-922-865-1	Sequence 1, Appli
c	35	37	7.4	1951	4	US-09-510-949-1	Sequence 1, Appli
	36	36.6	7.3	2136	4	US-09-252-991A-13089	Sequence 13089, A
c	37	36.6	7.3	2406	4	US-09-252-991A-12703	Sequence 12703, A
c	38	36.2	7.2	810	4	US-09-252-991A-3144	Sequence 3144, Ap
	39	36.2	7.2	1545	4	US-09-252-991A-2823	Sequence 2823, Ap
c	40	36	7.2	711	4	US-09-252-991A-12740	Sequence 12740, A
	41	36	7.2	804	4	US-09-252-991A-13365	Sequence 13365, A
c	42	36	7.2	1050	4	US-09-252-991A-12592	Sequence 12592, A
	43	36	7.2	1089	4	US-09-252-991A-13045	Sequence 13045, A
c	44	35.8	7.1	1506	4	US-09-252-991A-1159	Sequence 1159, Ap
c	45	35.8	7.1	2588	4	US-09-480-017-3	Sequence 3, Appli

```
                                ALIGNMENTS

RESULT 1
US-08-592-874-1/c
; Sequence 1, Application US/08592874
; Patent No. 5854034
; GENERAL INFORMATION:
;   APPLICANT:  POLLOCK, THOMAS J.
;   APPLICANT:  YAMAZAKI, MOTOHIDE
;   APPLICANT:  THORNE, LINDA
;   APPLICANT:  MIKOLAJCZAK, MARCIA
;   APPLICANT:  ARMENTROUT, RICHARD W.
;   TITLE OF INVENTION:  DNA SEGMENTS AND METHODS FOR INCREASING
;   TITLE OF INVENTION:  POLYSACCHARIDE PRODUCTION
;   NUMBER OF SEQUENCES:  1
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  JULES E. GOLDBERG
;     STREET:  261 MADISON AVENUE
;     CITY:  NEW YORK
;     STATE:  NY
;     COUNTRY:  USA
;     ZIP:  10016-2391
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/592,874
;     FILING DATE:
;     CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/377,440
;     FILING DATE:  24-JAN-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  GOLDBERG, JULES E.
;     REGISTRATION NUMBER:  24,408
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  212-986-4090
;     TELEFAX:  212-818-9479
; INFORMATION FOR SEQ ID NO:  1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  28804 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  unknown
;     TOPOLOGY:  unknown
;   MOLECULE TYPE:  DNA (genomic)
;   FRAGMENT TYPE:  N-terminal
US-08-592-874-1

  Query Match          17.7%;     Score 88.8;  DB 2;  Length 28804;
  Best Local Similarity 56.5%;    Pred. No. 2.7e-14;
```

Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```
Qy      80  TTCCTGCGCGCCGCCCTCGAAATTGTCGAGACGCCGCCATCTCCACCGGCGAGACTCACG 139
               |||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db    18808  TTCCTGCGCGCCGCCGCGCGCTCGAGATCATCGAACGTCCGGTGTCGCCCACCGCGCCTCACC 18749

Qy     140  GCCGCCTTGCTTGCTGCCTTGTTCTACTGGCGGGCGGGTCTCGGCAGGATC 199
               | | ||||||||||||||| |||| |||||||||||||||||||||||||
Db    18748  GGCCGGGTGATGGTGCGGCGGCTGGCGATCACCACCGCCTGGCAATCGGCCGCCGGTG 18689

Qy     200  GACATCGTTGCTTCTGCATCCAGAAAGATCGTGCCGGGCGACCGTGTAAAGCTGTTCAG 259
               ||| ||||||||||||| ||||  ||||||||||||| |||||||||||||||||
Db    18688  GAAGTGGTGCGGCGACGCCGCATCGGCGAGACCAAGATCGTGCAG 18629

Qy     260  CCGCTCGAGGTCGGCGTCGGTGCGGGCCACTCATGTCCGGGCGATGGCCAAACCGTCAAGGCC 319
               ||||| |||||||||||| |||||||||||||| |||||||||||||
Db    18628  TCCCCCGAAAGCGGTATCGTCCGGGCGCATTCTGGTGGGCGAGGGGCAGAAGGTTCAGAAG 18569

Qy     320  GGCGAGATTCTGATCGAGCTCGATCATCATTCGCGCGGGTGGGTGGATGTTGCGC 371
               |||| |||||||||||||| ||||||| | ||||
Db    18568  GGCCAGGTGCTGATCACGCTCGATCACCGCTCGATCCCACCATGTCGGCAGCCGAAGCCGCGGC 18517
```

RESULT 2
US-09-096-942-2/c
; Sequence 2, Application US/09099942
; Patent No. 6027925
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,942
; CURRENT FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 28804
;   TYPE: DNA
;   ORGANISM: Sphingomonas sp. S88
US-C9-096-942-2

Query Match   17.7%; Score 88.8; DB 3; Length 28804;
Best Local Similarity 56.5%; Pred. No. 2.7e-14;
Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```
Qy      80  TTCCTGCGCGCCGCCCTCGAAATTGTCGAGACGCCGCCATCTCCACCGGCGAGACTCACG 139
               |||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db    18808  TTCCTGCGCGCCGCCGCGCGCTCGAGATCATCGAACGTCCGGTGTCGCCCACCGCGCCTCACC 18749

Qy     140  GCCGCCTTGCTTGCTGCCTTGTTCTACTGGCGGGCGGGTCTCGGCAGGATC 199
               | | ||||||||||||||| |||| |||||||||||||||||||||||||
Db    18748  GGCCGGGTGATGGTGCGGCGGCTGGCGATCACCACCGCCTGGCAATCGGCCGCCGGTG 18689

Qy     200  GACATCGTTGCTTCTGCATCCAGAAAGATCGTGCCGGGCGACCGTGTAAAGCTGTTCAG 259
               ||| ||||||||||||| ||||  ||||||||||||| |||||||||||||||||
Db    18688  GAAGTGGTGCGGCGACGCCGCATCGGCGAGACCAAGATCGTGCAG 18629

Qy     260  CCGCTCGAGGTCGGCGTCGGTGCGGGCCACTCATGTCCGGGCGATGGCCAAACCGTCAAGGCC 319
               ||||| |||||||||||| |||||||||||||| |||||||||||||
Db    18628  TCCCCCGAAAGCGGTATCGTCCGGGCGCATTCTGGTGGGCGAGGGGCAGAAGGTTCAGAAG 18569

Qy     320  GGCGAGATTCTGATCGAGCTCGATCATCACGCTCGATCCCACCATGTCGCGCAGCCGAAGCCGCGGC 371
               |||| |||||||||||||| ||||||| || ||||
Db    18568  GGCCAGGTGCTGATCACGCTCGATCACCGCTCGATCCCACCATGTCGGCAGCCGAAGCCGCGGC 18517
```

RESULT 3
US-09-096-867-2/c
; Sequence 2, Application US/09096867
; Patent No. 6030817
; GENERAL INFORMATION:
; APPLICANT: Pollock, Thomas J
; APPLICANT: Mikolajczak, Marcia
; APPLICANT: Yamazaki, Motohide
; APPLICANT: Thorne, Linda
; APPLICANT: Armentrout, Richard W
; TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
; TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
; FILE REFERENCE: seq list for appl filed from pro. appl
; CURRENT APPLICATION NUMBER: US/09/096,867
; CURRENT FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/049,428
; EARLIER FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 28804
;   TYPE: DNA
;   ORGANISM: Sphingomonas sp. S88
US-C9-096-867-2

Query Match   17.7%; Score 88.8; DB 3; Length 28804;
Best Local Similarity 56.5%; Pred. No. 2.7e-14;
Matches 165; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```
Qy      80  TTCCTGCGCGCCGCCCTCGAAATTGTCGAGACGCCGCCATCTCCACCGGCGAGACTCACG 139
               |||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||
Db    18808  TTCCTGCGCGCCGCCGCGCGCTCGAGATCATCGAACGTCCGGTGTCGCCCACCGCGCCTCACC 18749

Qy     140  GCCGCCTTGCTTGCTGCCTTGTTCTACTGGCGGGCGGGTCTCGGCAGGATC 199
               | | ||||||||||||||| |||| |||||||||||||||||||||||||
Db    18748  GGCCGGGTGATGGTGCGGCGGCTGGCGATCACCACCGCCTGGCAATCGGCCGCCGGTG 18689

Qy     200  GACATCGTTGCTTCTGCATCCAGAAAGATCGTGCCGGGCGACCGTGTAAAGCTGTTCAG 259
               ||| ||||||||||||| ||||  ||||||||||||| |||||||||||||||||
Db    18688  GAAGTGGTGCGGCGACGCCGCATCGGCGAGACCAAGATCGTGCAG 18629

Qy     260  CCGCTCGAGGTCGGCGTCGGTGCGGGCCACTCATGTCCGGGCGATGGCCAAACCGTCAAGGCC 319
               ||||| |||||||||||| |||||||||||||| |||||||||||||
Db    18628  TCCCCCGAAAGCGGTATCGTCCGGGCGCATTCTGGTGGGCGAGGGGCAGAAGGTTCAGAAG 18569

Qy     320  GGCGAGATTCTGATCGAGCTCGATCATCACGCTCGATCCCACCATGTCGCGCAGCCGAAGCCGCGGC 371
               |||| |||||||||||||| ||||||| || ||||
Db    18568  GGCCAGGTGCTGATCACGCTCGATCACCGCTCGATCCCACCATGTCGGCAGCCGAAGCCGCGGC 18517
```

RESULT 4
US-08-772-270A-6
; Sequence 6, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
; APPLICANT: Ricciatti, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Bereskin & Parr
;   STREET: 40 King Street West
;   CITY: Toronto
;   STATE: Ontario
;   COUNTRY: Canada
;   ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
           GenCore version 5.1.6
     Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:15 ; Search time 2864 Seconds
                                         (without alignments)
                        1771.227 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 608
Sequence: 1 LSSLSPRMKSAREVVAVGGK........KAGEILIELDPFAGGVDVAT 124

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext 0.5
               Ygapop 10.0 , Ygapext 0.5
               Fgapop 6.0 , Fgapext 7.0
               Delop   6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn_1/USPTO_spool/USO9484577/runat_23102003_111410_17988/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=USO9484577 @CGN_1_1.3508 @runat 23102003_111410_17988 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQJERY -NEG_SCORES=0 -WAIT -DSPECOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
        1: gb_ba:*   29: em_vi:*
        2: gb_htg:*   30: em_htg_hum:*
        3: gb_in:*   31: em_htg_inv:*
        4: gb_om:*   32: em_htg_other:*
        5: gb_ov:*   33: em_htg_mus:*
        6: gb_pat:*   34: em_htg_pln:*
        7: gb_ph:*   35: em_htg_rod:*
        8: gb_pl:*   36: em_htg_mam:*
        9: gb_pri:*   37: em_htg_vrt:*
       10: gb_ro:*   38: em_sy:*
       11: gb_sts:*   39: em_htgo_hum:*
       12: gb_sy:*   40: em_htgo_mus:*
       13: gb_un:*   41: em_htgo_other:*
       14: gb_vi:*
       15: em_ba:*
       16: em_fun:*
       17: em_hum:*
       18: em_in:*
       19: em_mu:*
       20: em_om:*
       21: em_or:*
       22: em_ov:*
       23: em_pat:*
       24: em_ph:*
       25: em_pl:*
       26: em_ro:*
       27: em_sts:*
       28: em_un:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	312	51.3	302650	1	AP005958	Bradyrhiz
c	2	246	40.5	10902	1	AE011858	Xanthomon
	3	243	40.0	11180	1	AE013811	Yersinia
	4	243	40.0	334050	1	AJ414151	Yersinia
	5	241	39.6	15552	1	AE004049	Xylella f
	6	239	39.3	302022	1	AE012558	Xylella f
	7	238	39.1	1428	1	NME391263	Neisseria
	8	238	39.1	4254	1	AF121772	Neisseria
	9	238	39.1	11381	1	AE002524	Neisseria
c	10	238	39.1	349980	6	AX044033	Sequence
c	11	238	39.1	349980	6	AX044034	Sequence
	12	235	38.7	7824	1	NME391260	Neisseria
	13	235	38.7	326301	1	NMA622491	Neisseria
c	14	231.5	38.1	301708	1	AE016792	Pseudomon
c	15	218	35.9	189050	1	AL646077	Ralstonia
c	16	215	35.4	28804	1	SSU51197	Sphingomona
c	17	215	35.4	28804	6	AR068625	Sequence
	18	205	33.7	3526	1	AF273216	Rhizobium
	19	182	29.9	6952	1	AF314503	Mannheimi
	20	182	29.9	6952	1	AF314506	Mannheimi
	21	182	29.9	6952	1	AF314508	Mannheimi
	22	182	29.9	6952	1	AF314509	Mannheimi
	23	182	29.9	6952	1	AF314510	Mannheimi
	24	182	29.9	6952	1	AF314512	Mannheimi
	25	182	29.9	6952	1	AF314514	Mannheimi
	26	182	29.9	6952	1	AF314515	Mannheimi
	27	182	29.9	6952	1	AF414141	Mannheimi
	28	-182	29.9	6955	1	AF314511	Mannheimi
	29	182	29.9	6955	1	AF314513	Mannheimi
	30	182	29.9	6958	1	AF314523	Pasteurel
	31	182	29.9	6958	1	AF314524	Pasteurel
	32	182	29.9	6958	1	AF314526	Pasteurel
	33	182	29.9	7742	1	PASLKTCABD	Pasteurel
	34	181	29.8	6469	1	M24197	P.haemolyti
	35	181	29.8	6952	1	AF314521	Mannheimi
	36	181	29.8	6952	1	AF314517	Mannheimi
	37	181	29.8	6952	1	AF314518	Mannheimi
	38	181	29.8	6952	1	AF314519	Mannheimi
	39	181	29.8	6952	1	AF314520	Mannheimi
	40	181	29.8	6952	1	AF314522	Mannheimi
	41	181	29.8	300883	1	AE016809	Vibrio vu
	42	179	29.4	7801	1	PASALKT	Pasteurella
	43	179	29.4	7841	1	PAJLEUKOT	Pasteurella
	44	177	29.1	7489	1	PAU66588	Pasteurella
	45	176	28.9	15746	1	APXIIICAB	A.pleuropne

ALIGNMENTS

RESULT 1

```
AP005958/c
LOCUS       AP005958      302650 bp    DNA    linear   BCT 28-MAR-2003
DEFINITION  Bradyrhizobium japonicum USDA 110 DNA, complete genome, section
            24/31.
ACCESSION   AP005958 BA000040
VERSION     AP005958.1  GI:27354550
KEYWORDS    .
SOURCE      Bradyrhizobium japonicum USDA 110
  ORGANISM  Bradyrhizobium japonicum USDA 110
            Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
            Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE   1
  AUTHORS   Kaneko,T., Nakamura,Y., Sato,S., Minamisawa,K., Uchiumi,T.,
            Sasamoto,S., Watanabe,A., Idesawa,K., Iriguchi,M., Kawashima,K.,
            Kohara,M., Matsumoto,M., Shimpo,S., Tsuruoka,H., Wada,T., Yamada,M.
            and Tabata,S.
  TITLE     Complete genomic sequence of nitrogen-fixing symbiotic bacterium
            Bradyrhizobium japonicum USDA110
  JOURNAL   DNA Res. 9 (6), 189-197 (2002)
   MEDLINE  22484998
   PUBMED   12597275
REFERENCE   2
  AUTHORS   Kaneko,T., Nakamura,Y., Sato,S., Minamisawa,K., Uchiumi,T.,
            Sasamoto,S., Watanabe,A., Idesawa,K., Iriguchi,M., Kawashima,K.,
            Kohara,M., Matsumoto,M., Shimpo,S., Tsuruoka,H., Wada,T., Yamada,M.
            and Tabata,S.
  TITLE     Complete genomic sequence of nitrogen-fixing symbiotic bacterium
            Bradyrhizobium japonicum USDA110 (supplement)
  JOURNAL   DNA Res. 9 (6), 225-256 (2002)
   MEDLINE  22485002
   PUBMED   12597279
REFERENCE   3  (bases 1 to 302650)
  AUTHORS   Kaneko,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Takakazu Kaneko, Kazusa DNA Research
            Institute, The First Laboratory for Plant Gene Research; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:kaneko@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/rhizobase/,
            Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1..302650
                     /organism="Bradyrhizobium japonicum USDA 110"
                     /mol_type="genomic DNA"
                     /strain="USDA110"
                     /db_xref="taxon:224911"
     gene            31..1110
                     /gene="blr6272"
     CDS             31..1110
                     /gene="blr6272"
                     /note="AraC family
                     ORF_ID:blr6272"
                     /codon_start=1
                     /transl_table=11
                     /product="transcriptional regulatory protein"
                     /protein_id="BAC51537.1"
                     /db_xref="GI:27354551"
                     /translation="MENLLESLVESFGLGGLPVAASGRDAKRGRAMFTDLQAAPAWID
                     LQAEPRDHQLRQAAQAIPRERGWRRIDRGPEAAIEDITISRWEDFRDSSHDATTPED
                     RYFVGIALKATRAKLTRDRQVIFDGTMPAGTLYVSAPSKPLAAQFQAPCAFLHFHISA
                     DHFPAQSPAAEGLGDLVLLRDPLAAELARALIEHGDAADREFTRCVGQTLAMHARLE
                     FPRARVNALPKWRLRRVEQYIADNFDRCISLSELANVAGLSHHFAQFRTATGYRPR
                     EYLLNHRIEHAKTLLATTGRPLAEIALAVGFSTQAHFSTVFKRISGQSPARWRLASKA
                     EPLAEAALPRRPAADRDWMASAAA"
     gene            complement(1428..3794)
                     /gene="bll6273"
     CDS             complement(1428..3794)
                     /gene="bll6273"
                     /note="ORF_ID:bll6273
                     probable biotin sulfoxide reductase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51538.1"
                     /db_xref="GI:27354552"
                     /translation="MDRYTATMYRNKDQMMDDTIGFDPDPGLLLSDGFKPHTSHWGVFS
                     ARNGAAGLEVRAYAGDPDPNGIIDNFPGALRHQARIAQPAIRRGWLERGPGDDRRGR
                     DEFVSVSWEKALDLLGDELGRIRDTRGPGAVFGGSYGWSSAGRFHHAQSQVHRFLNIA
                     MGGYVRSVNSYSSGASSVLLPQILAGYEDITKRNVTWEQIATETDIVLAFGMALKNS
                     MVAGGSISKHVERSAMEAARKRGCEFIIVSPLREDLPVEAGAEWMTCVPSTDTALMLG
                     IVHTIVSEKIHQAFLDRYTEGWPVFLRYLTGBSDGQPKSAEWAAAISGVDADTIRTL
                     ARRLAGKRALITVSHLSQRAEHGEQPVWMGMVLAAALGQIGLPGGGYAYSLGAIGYYG
                     RRVNDVPGTLGQGRNGVRDFIPVARIADMLLNPGATYRYNGETRTYPDIRLVYWAGG
                     NPFHHQDINRLRKAFAQLDTFVVHELAWTATARHADIVLPSTMTLBREDIGYSSNDP
                     LMVAMHRIAEPFGLARDDYEIFADLAERLGAREPFTEGRTSREWLKHLYEPTRASLAM
                     RGLEAPSFDEFWQRGSLVVPQQPDDGGRLRRFREDVAHALPTPSGRIEIFSTRIAGH
                     GDADCPGHVWLDKTDMPKPGAPCFLVANQPVTRLHSQLDFGGHSLGAKHRGREVARM
                     NPVDADARGIKDGDIIRLFNERGACLAAVHVTDGIAPGVVQLPTGAWYDPMDPEDDAP
                     LCVHGNPNVLTRDVGTSSFAQGCTGQLTTVEVEKFTGNLPPIRAFDPV"
     gene            complement(3898..4947)
                     /gene="bll6274"
     CDS             complement(3898..4947)
                     /gene="bll6274"
                     /note="ORF_ID:bll6274
                     putative dehydrogenase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51539.1"
                     /db_xref="GI:27354553"
                     /translation="MTGASDTVAIGADALRPFAADLLRAGGFTADQADETAAMLVWAN
                     LRGADSHGVLRIPRYVEMVELGLINPAAEPRQVSGKGAVAVIDAQRAPGAVGMSLAAR
                     KASELADTHGIGWCAVRSISHAGAIGYVAQAGRIGIVMTASKPLMIYHGSSAEG
                     VSTNPIAIAAPTGNPDRPLLLDMSTASVALGKIMAARDAGQPIPVGWAVDGNGQATTD
                     PREAKAVLPMAGPKGSGLSLMIEVLASVLTANPLIAVALVHGGDPGGNGLVAVLDFSA
                     FGDNFTLEVDRLGAAIKSLPPAAGADSVYLPGERGFEMERRSRAGIPLARGTVSRL
                     ALASKLNVAVPEGVA"
     gene            complement(4950..5873)
                     /gene="bll6275"
     CDS             complement(4950..5873)
                     /gene="bll6275"
                     /note="ORF_ID:bll6275
                     hypothetical protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51540.1"
                     /db_xref="GI:27354554"
                     /translation="MSAPPLSQPTREIPMQEKHTFRSDGLKLSAVLHVPDARKPGQRL
                     PAFIVLHGFVGSKDESHABIQARMLEQMGYVALRFDFRCCGBSEGERAQVRCFDQVAD
                     AKNALTFLAEREVDPGRIGVVGHSGAAVSYVAAGVDDRFACVISSCGWGDGERKPR
                     GQHPTPEAWAKFTSLLEKGRKQKQETCGQSMWISRFDAVPIFEHLRKNLSPKAIMEIPV
                     ETAWSMYNFRADEVVGNIAPRPLFLHTANDTITPTEQSIRMFEKAGQPAELVLITGT
                     SHFPLAEGDAPRTKAIIKGWLDKFFPSPLGA"
     gene            complement(5910..6812)
                     /gene="bll6276"
     CDS             complement(5910..6812)
                     /gene="bll6276"
                     /note="ORF_ID:bll6276
                     unknown protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAC51541.1"
                     /db_xref="GI:27354555"
                     /translation="MPPIVHAVGHVKINTTAIEAVVRDATEILGLQITYTDEGQTWLS
                     SNGRAAELVLLRSHENATHTIGLEALTVDAVREAESRVESAGCRVISRKPSLDCMAAG
                     VTFDTPEGLRFEVHTPTRDQIVGRRYPTHGVGPNRMDHLNLTSPDPIATRQQLERICG
                     LRISERMVDDSLSMMYAGGNRQHHIVGVVRGKVGLHHYSFBPKEFNDYLKLGDLLRFD
                     RQMLWGPGRHRPGDNTYAYTDSSGAMIECSGSMALIADDDGFVPNVITNLERPGNVR
                     CMNVWGTPAPLEWRQFHPFTAIT"
     gene            7080..7853
                     /gene="blr6277"
     CDS             7080..7853
                     /gene="blr6277"
                     /note="GntR family
                     ORF_ID:blr6277"
                     /codon_start=1
                     /transl_table=11
                     /product="transcriptional regulatory protein"
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:20 ; Search time 1724 Seconds
(without alignments)
1748.118 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 608
Sequence: 1 LSSLSPRMKSAREVAVGGK.........KAGEILIELDPFAGGVDVAT 124

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 1215223856 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09484577/runat_23102003_111410_18001/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -L=ST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09484577 @CGN_1_1_2810 @runat_23102003_111410_18001 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
C	1	122	20.1	772	28	AF075872	AF075872 AF075872
C	2	117	19.2	1126	29	BZ573090	BZ573090 msh2_2939
C	3	105	17.3	750	28	AF075952	AF075952 AF075952
C	4	95.5	15.7	945	10	BE564254	BE564254 601343055
	5	92	15.1	812	29	BZ549462	BZ549462 pacs1-60
	6	87	14.3	961	10	BF781847	BF781847 602107387
	7	85.5	14.1	455	29	P284R	P284R Leishmani
	8	85	14.0	682	29	AG163349	AG163349 Pan trogl
	9	84	13.8	432	10	BG274115	BG274115 WHE2231_G
C	10	84	13.8	756	28	AQ991716	AQ991716 Rfc00179F
	11	83.5	13.7	610	12	BM631476	BM631476 170006875
	12	83.5	13.7	973	10	BG758850	BG758850 602711239
C	13	83	13.7	939	29	BZ706228	BZ706228 PUBLV37TD
	14	82.5	13.6	431	29	BZ883003	BZ883003 CH240_241
C	15	82.5	13.6	1175	29	BZ558366	BZ558366 pacs1-60
	16	82	13.5	456	12	BJ189099	BJ189099 BJ189099_
C	17	82	13.5	884	29	BZ636906	BZ636906 OGCCK65TM
C	18	81.5	13.4	563	28	BH775270	BH775270 fzmb011f0
	19	81.5	13.4	678	14	CD425333	CD425333 SA1_11_H0
C	20	81	13.3	469	9	AU070712	AU070712 AU070712
	21	81	13.3	906	14	CB575007	CB575007 AGENCOURT
	22	80	13.2	795	29	CNS05PMK	CNS05PMK Tetraodon
C	23	80	13.2	1201	9	AL514080	AL514080 AL514080
	24	80	13.2	1337	29	BZ572952	BZ572952 msh2_2874
C	25	79.5	13.1	360	10	BE481913	BE481913 167524 BA
	26	79.5	13.1	394	14	D27997	D27997 CELK007B1F
	27	79.5	13.1	423	9	AI969323	AI969323 wq96h04.x
C	28	79.5	13.1	507	13	BQ606908	BQ606908 BRY_2784
C	29	79.5	13.1	532	13	BQ817896	BQ817896 103006 6H0
C	30	79.5	13.1	768	12	BI909889	BI909889 603070521
C	31	79.5	13.1	1002	13	BX344531	BX344531 BX344531
	32	79.5	13.1	1088	29	BZ577484	BZ577484 msh2_5430
	33	79	13.0	487	10	BG158865	BG158865 RHI22_43
C	34	79	13.0	710	12	BJ570919	BJ570919 BJ570919
C	35	79	13.0	711	12	BJ568037	BJ568037 BJ568037
C	36	79	13.0	946	10	BF305164	BF305164 601892608
	37	79	13.0	969	10	BF787122	BF787122 602108661
	38	79	13.0	1159	29	AG071759	AG071759 Pan trogl
	39	78.5	12.9	499	14	CA621031	CA621031 wlin.pk00
C	40	78.5	12.9	528	14	CB993505	CB993505 AGENCOURT
C	41	78.5	12.9	734	28	AQ272038	AQ272038 nbxb00270
C	42	78.5	12.9	775	29	BZ577887	BZ577887 msh2_5621
	43	78.5	12.9	825	12	BI603588	BI603588 603244491
	44	78.5	12.9	852	29	BZ553772	BZ553772 pacs1-60
	45	78.5	12.9	853	29	CC115510	CC115510 NDL.7lC2_

ALIGNMENTS

RESULT 1
AF075872/c
LOCUS   AF075872   772 bp   DNA   linear   GSS 29-AUG-2000
DEFINITION AF075872 Salmonella typhimurium LT2, Lambda DASH II Salmonella
            typhimurium genomic clone 142-T3, genomic survey sequence.
ACCESSION   AF075872
VERSION   AF075872.1 GI:3320742
KEYWORDS   GSS.
SOURCE   Salmonella typhimurium
ORGANISM   Salmonella typhimurium
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Salmonella.
REFERENCE   1 (bases 1 to 772)

AUTHORS   Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.
TITLE   Sample sequencing of a Salmonella typhimurium LT2 lambda library:
             comparison to the Escherichia coli K12 genome
JOURNAL   FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE   99243757
PUBMED   10227170
COMMENT   Contact: McClelland M
             Molecular Biology
             Sidney Kimmel Cancer Center
             3099 Science Park Road, San Diego, CA 92121, USA
             Email: mcclelland@lifsci.sdsu.edu
             Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..772
                     /organism="Salmonella typhimurium"
                     /mol_type="genomic DNA"
                     /strain="LT2"
                     /db_xref="taxon:602"
                     /clone="142-T3"
                     /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
                     /note="Vector: Lambda DASH II; sequenced using Li-Cor
                     sequencer"
BASE COUNT   169 a   229 c   205 g   167 t   2 others
ORIGIN

Alignment Scores:
Pred. No.:   0.00187
Score: 122.00   Length: 772
Percent Similarity: 61.11%   Matches: 26
Best Local Similarity: 36.11%   Conservative: 18
Query Match: 20.07%   Mismatches: 28
DB: 28   Indels: 1
                                               Gaps: 0

US-09-484-577A-4 (1-124) x AF075872 (1-772)

Qy   49 LeuLeuAlaAlaLeuPheTyrCysAlaValAlaThrAlaValAlaTrpAlaGlyLeuGlyArgIleAspIle 68
         ::: ||||| |||| |||| | |||:||::|| ||| ::|||
Db   336 ATCTGCTCACTGTTATTTCTCATTCTGGCCATCTGGGCGTGGTTGGCCGACTGGATGGATGAG 277

Qy   69 ValAlaSerAlaSerArgGlysIleValProGlyAspArgValLysLeuValGlnProLeu 88
         ||:|| :: ||||| ||||:||| :::|||||| |||
Db   276 GTTTCCACCGGACAACGGAAGTGATCCCCAGTTCACGCGAAGCAGGTTCTGCAGTCGCTG 217

Qy   89 GluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGlyGlu 108
         ::: |||:|||||| |||| ||||||||| ||||||||||||
Db   216 GATGGCGGCATTCTGGCGAGTTGACGGTGCGGAAGGCGACAGAGTTCAGGTAACCAG 157

Qy   109 IleLeuIleGluLeuAspProPheAlaGlyGlyVal 120
         ||::::: |||||||| ||||| ||||||
Db   156 ATTGTCGCCCCGGCTTGATCCGAC-GCGTCTGGCGTC 122

RESULT 2
BZ573090/c
LOCUS   BZ573090   1126 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION   msh2_2939.y2 msh Pseudomonas aeruginosa genomic clone mst2_2939,
             genomic survey sequence.
ACCESSION   BZ573090
VERSION   BZ573090.1 GI:27208151
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
  ORGANISM   Pseudomonas aeruginosa
             Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
             Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 1126)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Smith,E.E., Hastings,M.,
             Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE   Whole-Genome-Sequence variation among multiple isolates of
             Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
  COMMENT   Contact: Chris K. Raymond
             Genome Center
             University of Washington
             Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..1126
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="MSH"
                     /db_xref="taxon:287"
                     /clone="msh2_2939"
                     /clone_lib="msh"
                     /note="Environmental isolate. Whole genomic shotgun
                     library."
BASE COUNT   170 a   360 c   314 g   282 t
ORIGIN

Alignment Scores:
Pred. No.:   0.0107
Score: 117.00   Length: 1126
Percent Similarity: 47.58%   Matches: 38
Best Local Similarity: 30.65%   Conservative: 21
Query Match: 19.24%   Mismatches: 51
DB: 29   Indels: 14
                                               Gaps: 5

US-09-484-577A-4 (1-124) x BZ573090 (1-1126)

Qy   3 SerLeuSerProArgMetLysSer-----AlaArgGluValValAlaValAlaValGlyGlyLys 20
         ||| ::||||||| ||| ::: ||| ||||:|
Db   734 TCATCGACGCCCCCGTCGAAGCAGTGGTCAGCAGTTCAGGGTCCAGGCCCGC 675

Qy   21 ThrArgAspGlu--------LeuAlaPheLeuProAlaAlaLeuGluValIleValGluThr 37
         ::: ||||| ||||:||| ||||||||||||||
Db   674 CGAAGGAGCCAACCATGATCTCTGACCGCAAACCCCGCCCGCCCGGCAACTGACC 615

Qy   38 ProProSerProThrAlaArgLeuThrAlaAlaLeuLeuAlaAlaLeuPheTyr----- 55
         |||||||| ::: ||| ||||||
Db   614 GACCCGTTGCTGACGGCT-----ACCCACCCGGTCTACCGGCGTGTTCTGGACCTG 561

Qy   56 -------CysAla--------ValAlaTrpAlaGlyLeuGlyArgIleAspIleValAla 70
         |||| :::||| ||||| ||:|||
Db   560 CTCGGTTGCCTGTCTGTCTCATCGGCGGCCGCCAACTGGACCAGAGGTGACC 501

Qy   71 SerAlaSerArgLysIleValProGlyAspArgValValLysLeuValGlnProLeuGlyVal 90
         |||::::||| |||||||| |||| :: |||||
Db   500 CGCGGCGACGGTCGGGTGCGTCCGTTCAGCCGCATCCAGACAGCTGGAGGGC 441

Qy   91 GlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGlyLeuLeu 110
         ||:::::   :::|||||| ||||| ||:||| :::|| |||
Db   440 GGCATCCTCGACCCGCCTCGGTGAAGGAGGGCGACCTGGTGAAGTCGGCCAGCCGCGCTG 381

Qy   111 IleGluLeuAsp 114
         ::: |||||||
Db   380 GTGCGCCTCGAC 369

RESULT 3
AF075952
LOCUS   AF075952   750 bp   DNA   linear   GSS 29-AUG-2000
DEFINITION   Salmonella typhimurium LT2, Lambda DASH II Salmonella
             typhimurium genomic clone 1256-T3, genomic survey sequence.
ACCESSION   AF075952
VERSION   AF075952.1 GI:3320822
KEYWORDS   GSS.
SOURCE   Salmonella typhimurium
  ORGANISM   Salmonella typhimurium
             Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
             Enterobacteriaceae; Salmonella.
REFERENCE   1 (bases 1 to 750)
  AUTHORS   Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.
  TITLE   Sample sequencing of a Salmonella typhimurium LT2 lambda library:
             comparison to the Escherichia coli K12 genome
  JOURNAL   FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
  MEDLINE   99243757

OM protein - nucleic search, using frame_plus_p2n model

Copyright (c) 1993 - 2003 Compugen Ltd.
GenCore version 5.1.6

Run on: October 23, 2003, 11:14:14 ; Search time 236 Seconds
(without alignments)
1418.349 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 608
Sequence: 1 LSSLSPRMKSAREVVAVGGK.......KAGEILIELDPFAGGVDVAT 124

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model --DEV=xlh
-Q=/cgn2.1/USPTO_spool/US0948577//runat_23102003_111409_17977/app_query.fasta_1.263
-DB=N Geneseq_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0948577 @CGN_1_1_0 @runat_23102003_111409_17977 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -YGAPEXT=7 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	238	39.1	1428	21	AAZ54335	Neisseria meningit
	2	238	39.1	24158	21	AAA81532	N. meningitidis pa
c	3	238	39.1	349980	21	AAF21611	Neisseria meningit
c	4	238	39.1	349980	21	AAF21612	Neisseria meningit
c	5	238	38.7	837096	21	AAA81489	N. meningitidis pa
	6	235	35.4	1305	24	AAZ54336	Neisseria meningit
c	7	215	35.4	28804	17	AAT37329	Sphingan biosynthe
	8	215	35.4	28804	18	AAT92474	Sphingomonas genus
c	9	215	35.4	28804	20	AAV99812	Sphingomonas S88 s
c	10	215	35.4	28804	20	AAV81474	Chromosomal fragme
	11	190	31.2	7184	15	AAQ70050	ptx gene of Pasteu
	12	182	29.9	7183	11	AAQ06074	Sequence encoding
	13	176	28.9	76804	24	ABS78942	E. coli CFT073 gen
	14	169	27.8	5120	18	AAT73218	ApxIBD gene. Acti
	15	169	27.8	5120	21	AAZ88585	A. pleuropneumonia
	16	168	27.6	7721	18	AAT73220	ApxIIABCD gene.
	17	168	27.6	7721	21	AAZ88587	A. pleuropneumonia
c	18	154	25.3	3576	19	AAV31265	E. coli J96 pathog
	19	124.5	20.5	1245	22	ABA89180	Escherichia coli p
	20	124.5	20.5	48012	24	ABS79084	E. coli CFT073 gen
	21	124.5	20.5	48254	22	ABA89141	Escherichia coli p
	22	124.5	20.5	48345	22	ABA89142	Escherichia coli p
	23	104	17.1	6400	20	AAZ22701	Pseudomonas fluore
	24	99	16.3	4466	17	AAT39662	Esterase secretory
c	25	91	15.0	2382	24	AAS94828	Human DNA sequence
c	26	90.5	14.9	9320	23	AAH45588	DNA encoding hydro
	27	90	14.8	1302	23	ABL41686	ABC transporter T1
	28	90	14.8	8580	23	ABL41682	ABC transporter ge
	29	89	14.6	5660	24	AAZ22700	Pseudomonas fluore
c	30	88	14.5	834	23	ABL26523	Drosophila melanog
c	31	88	14.5	2834	23	ABL26522	Drosophila melanog
	32	88	14.5	13518	20	AAX20563	Polynucleotide seq
c	33	87.5	14.4	4403765	22	AAI99683	Mycobacterium tube
c	34	87.5	14.4	441529	23	AAI99682	Mycobacterium tube
c	35	86	14.1	11308	23	AAS59625	Propionibacterium
	36	81.5	13.4	1395	21	AAA65929	E. coli proliferat
	37	81.5	13.4	1395	22	AAH84642	E. coli growth and
	38	81.5	13.4	2310	22	AAS82449	DNA encoding novel
	39	80.5	13.2	3462	22	AAS01509	Bacillus subtilis
c	40	80.5	13.2	4540	24	ABK63777	Rat sequence diffe
c	41	79	13.0	19053	21	AAA59146	Nucleotide sequenc
	42	78.5	12.9	357	21	AAC75328	Human ORFX ORF883
	43	78.5	12.9	357	21	ABN21772	Human ORFX polynuc
c	44	78	12.8	6848	22	AAK83993	Human immune/haema
c	45	77.5	12.7	2556	25	ABZ40207	N. gonorrhoeae nuc

ALIGNMENTS

RESULT 1
AA254335
ID   AA254335 standard; DNA; 1428 BP.
XX
AC   AAZ54335;
XX
DT   21-MAR-2000 (first entry)
XX
DE   Neisseria meningitidis ORF 764 partial DNA sequence SEQ ID NO:2619.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.

```
Db        334 GACGGGCAGCATGTGAAACAGGGAGAAACGTGGCCGGAACTGGAGGCTGTGGGAACAGAC 393

Qy        120 ValAspVal 122
                  ||||||
Db        394 AGCGGATGTG 402

RESULT 2
AAA81532
ID   AAA81532 standard; DNA; 24158 BP.
XX
AC   AAA81532;
XX
DT   04-DEC-2000  (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_79 SEQ ID NO:79.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200022430-A2.
XX
PD   20-APR-2000.
XX
PF   99WO-US23573.
XX
PR   08-OCT-1999;   98US-0103794.
PR   09-OCT-1998;   98US-0103794.
PR   30-APR-1999;   99US-0132068.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Frazer CM, Hickey E,  Peterson J,  Tettelin H, Venter JC;
PI   Masignani V, Galeotti C,  Mora M,  Ratti G,  Scarselli M,  Scarlato V;
PI   Rappuoli R,  Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be
PT   used in the diagnosis and treatment of N. meningitidis infection and
PT   other Neisserial infections, for example, N.gonorrhoea -
XX
PS   Claim 7; Page 1464-1471; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic
CC   proteins from Neisseria genomic sequences. AAA81453 to AAA82214
CC   represent specifically claimed Neisseria meningitidis genomic DNA
CC   sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC   Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC   AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC   isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC   AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC   sequences, which are all used in the exemplification of the present
CC   invention. The nucleic acid sequences, protein sequences, and antibodies
CC   against them, can be used in the manufacture of a composition. The
CC   composition can be used as a medicament (or in the manufacture of a
CC   medicament) for treating, preventing or diagnosing infection due to
CC   Neisserial bacteria. For example, some of the identified proteins could
CC   be components of vaccines against Meningococcus B; against all serotypes;
CC   and/or against all pathogenic Neissariae. Identification of sequences
CC   from the bacterium will also facilitate production of biological probes,
CC   particularly organism-specific probes. Attempts to make efficacious
CC   Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC   Multivalent vaccines have also been tried but none have successfully
CC   overcome antigenic variability. The provision of further, complete
CC   sequences may provide an opportunity to identify secreted or surface
CC   exposed proteins that may be presumed targets for the immune system and
CC   which are not antigenically variable or at least more conserved than
CC   other more variable regions.
XX
SQ   Sequence 24158 BP; 5559 A; 6297 C; 6241 G; 6061 T; 0 other,
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:     October 23, 2003, 12:11:35 ; Search time 234 Seconds
                                          (without alignments)
                                          1421.125 Million cell updates/sec

Title:          US-09-484-577A-4
Perfect score:  608
Sequence:       1 LSSLSPRMKSAREVVAVGGK.........KAGEILIELDPFAGGVDVAT 124

Scoring table:  BLOSUM62
                Xgapop   10.0 ,  Xgapext  0.5
                Ygapop   10.0 ,  Ygapext  0.5
                Fgapop    6.0 ,  Fgapext  7.0
                Delop     6.0 ,  Delext   7.0

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Searched:       1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters:       3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Command line parameters:
MODEL=frame+_p2n.model -DEV=xlh
Q=/cgn2.1/USPTO spool/US09484577/runat 23102003 111412 18082/app query.fasta_1.263
DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
MAXLEN=2000000000 -USER=US09484577 @CGN 1_285 @runat_23102003_111412_18082
NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

No.						DB ID	Description
2		176	28.9	76804	12	US-10-085-959-109	Sequence 109, App
3	c	154	25.3	3576	9	US-09-956-004-79	Sequence 79, Appl
4		124.5	20.5	1245	12	US-10-238-075-906	Sequence 906, App
5		124.5	20.5	48012	12	US-10-085-959-251	Sequence 251, App
6		124.5	20.5	48254	12	US-10-238-075-829	Sequence 829, App
7		109	17.9	1188	12	US-10-246-330-9	Sequence 9, Appli
8		103	16.9	25165	14	US-10-114-170-39	Sequence 39, Appl
9	c	91	15.0	2382	12	US-10-240-965-83	Sequence 83, Appl
10		84	13.8	1389	14	US-10-156-761-6704	Sequence 6704, Ap
11		84	13.8	9025608	14	US-10-156-761-1	Sequence 1, Appli
12		81.5	13.4	1395	9	US-09-912-020-121	Sequence 121, App
13		81.5	13.4	1395	12	US-10-287-274-270	Sequence 270, App
14	c	80.5	13.2	4540	10	US-10-917-800A-1684	Sequence 1684, Ap
15	c	78	12.8	2398	14	US-10-114-170-260	Sequence 260, Ap
16		77	12.7	11058	14	US-10-156-761-3629	Sequence 3629, Ap
17	c	77	12.7	9025608	14	US-10-156-761-1	Sequence 1, Appli
18		76.5	12.6	495	14	US-10-156-761-3641	Sequence 3641, Ap
19	c	76.5	12.6	1660	9	US-09-734-569-145	Sequence 145, App
20		75.5	12.4	3513	14	US-10-156-761-6676	Sequence 6676, Ap
21	c	75	12.3	2177	12	US-10-247-671-97	Sequence 97, Appl
22		75	12.3	2196	14	US-10-084-817-120	Sequence 120, App
23		74.5	12.3	2307	11	US-09-893-519A-87	Sequence 87, Appl
24	c	74	12.2	1473	14	US-10-156-761-4268	Sequence 4268, Ap
25		73.5	12.1	234	10	US-09-974-300-4611	Sequence 4611, Ap
26	c	73.5	12.1	2715	14	US-10-156-761-2245	Sequence 2245, Ap
27		73	12.0	585	12	US-10-029-386-1336	Sequence 1336, Ap
28		73	12.0	714	13	US-10-027-632-135196	Sequence 135196,
29		73	12.0	45191	15	US-10-080-170-649	Sequence 649, Ap
30	c	72.5	11.9	984	14	US-10-156-761-2093	Sequence 2093, Ap
31	c	72.5	11.9	1188	9	US-09-815-242-9794	Sequence 9794, Ap
32		72.5	11.9	2469	14	US-10-156-761-3389	Sequence 3389, Ap
33	c	72.5	11.9	3507	10	US-09-938-842A-105	Sequence 105, App
34		72	11.8	1308	9	US-09-730-772-12	Sequence 12, Appl
35		72	11.8	1308	9	US-09-735-849-12	Sequence 12, Appl
36		72	11.8	1308	12	US-10-379-830-12	Sequence 12, Appi
37		72	11.8	1308	12	US-09-574-819-12	Sequence 12, Appl
38	c	72	11.8	1578	14	US-10-156-761-4119	Sequence 4119, Ap
39		72	11.8	3668	12	US-10-204-041-13	Sequence 13, Appl
40	c	72	11.8	3990	9	US-09-815-242-6015	Sequence 6015, Ap
41		71.5	11.8	32767	12	US-10-004-113-4	Sequence 4, Appli
42		71.5	11.8	3222	9	US-09-815-242-8263	Sequence 8263, Ap
43		71.5	11.8	3441	9	US-09-815-242-4413	Sequence 4413, Ap
44		71.5	11.8	5030	8	US-08-781-986A-324	Sequence 324, App
45	c	71	11.7	559	12	US-10-029-386-6586	Sequence 6586, Ap

ALIGNMENTS

```
RESULT 1
US-09-884-696-36
; Sequence 36, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT: ANGELOS, JOHN A
;  APPLICANT: HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION: BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE: 2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
;  LENGTH: 1428
;  TYPE: DNA
;  ORGANISM: Moraxella bovis
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(1425)
US-09-884-696-36
```

```
Published Applications NA:*
 1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
-----	-----	-----	------	-----	-----------	-----------
1	242	39.8	1428	11	US-09-884-696-36	Sequence 36, Appl

Alignment Scores:
Pred. No.:   2.37e-23   Length: 1428
Score: 242.00   Matches: 50
Percent Similarity: 65.79%   Conservative: 25
Best Local Similarity: 43.86% Mismatches: 39
Query Match: 39.80%   Indels: 0
DB: Gaps: 0

US-09-484-577A-4 (1-124) x US-09-884-696-36 (1-1428)

```
Qy    10  SerAlaArgGluValValAlaValGlyGlyLysThrArgAspArgGluLeuAlaPheLeuPro  29
              :::  ||||  ::   ::||:::||:|||||||||||||
Db    64  GCAGTTCGAGACCAACTAACCCCTCCTAAGCGTACTAAAGAAGAACTCGCTTTTCTTCCT  123

Qy    30  AlaAlaLeuGluIleValGluThrProSerProThrAlaArgLeuThrAlaAlaLeu  49
              ::||||||  :::  |||||||  :
Db   124  GCACATCTAGAACTCACTGACACACCTGTATCCAGATCCTTCTAAGTGGACAGCTAGAATA  183

Qy    50  LeuAlaAlaAlaLeuPheTyrCysAlaValAlaAlaTrpAlaValAlaGlyLeuGlyArgIleAspIleVal  69
              ::  |||:::   :::||:|:||||:|:|||||
Db   184  ATCATGATATTTGTCCTATTTGCTTGCTATGGTCTTCTGGGTTGGACAGATTGACATTGTT  243

Qy    70  AlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGlnProLeuGlu  89
              ||:::||||||||  ||||  |||  :||||  ::|:  |||||
Db   244  GCTACAGCTTCAGGTAAAATTTCTTCAGGTAGCCGTAGCCGTAGCAAGACTATTCAATCTTTGGA  303

Qy    90  ValGlyValValArgGlyAlaThrHisValArgAspGlyGlnThrValLysAlaGlyGluIle  109
              :::|||||||||  :::|||  |||  ||||||||
Db   304  ACAGCGATAGTTAAAGCAGTTTATGTACGTGATGGTCAAAATGTTCAACAAGGTGAAATA  363

Qy   110  LeuIleGluLeuAspProPheAlaGlyValGlyValAspValAla  123
              ||:::::||||  |||||||
Db   364  TTAGTAGATTTAGTGGGAATCGGTTCAGATAGTGATGTTGCT  405
```

RESULT 2
US-10-085-959-109
, Sequence 109, Application US/10085959
, Publication No. US20030165870A1
GENERAL INFORMATION:
, APPLICANT: Blattner, Frederick R.
, APPLICANT: Welch, Rodney A.
, APPLICANT: Burland, Valerie D.
TITLE OF INVENTION: No. US2003016587OA1e1 Sequence of E. Coli CFT073
FILE REFERENCE: 960296.97648
CURRENT APPLICATION NUMBER: US/10/085,959
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: 60/242,412
PRIOR FILING DATE: 2000-10-19
NUMBER OF SEQ ID NOS: 255
SOFTWARE: PatentIn version 3.1
SEQ ID NO 109
LENGTH: 76804
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (39751)..(39751)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (49372)..(49372)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (66539)..(66539)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73693)..(73693)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73705)..(73705)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73737)..(73737)
OTHER INFORMATION: Unsure

NAME/KEY: misc_feature
LOCATION: (73759)..(73759)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73761)..(73761)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73766)..(73766)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73774)..(73774)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73781)..(73781)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73798)..(73798)
OTHER INFORMATION: Unsure
NAME/KEY: misc_feature
LOCATION: (73802)..(73802)
OTHER INFORMATION: Unsure
US-10-085-959-109

Alignment Scores:
Pred. No.:   6.11e-12   Length: 76804
Score: 176.00   Matches: 36
Percent Similarity: 60.78%   Conservative: 26
Best Local Similarity: 35.29% Mismatches: 40
Query Match: 28.95%   Indels: 0
DB: Gaps: 0

US-09-484-577A-4 (1-124) x US-10-085-959-109 (1-76804)

```
Qy    20  LysThrArgAspGluValAlaPheLeuAlaLeuProAlaAlaAlaGluIleValGluThrProPro  39
              :::   ::|:||||||||||||  ||||||||::|||||||||
Db 17983  CGTGAAAAGGACGAAAAATGAATTCTTACCCGCTCATCCGGAATTAATTGAAACGCCGGTA  18042

Qy    40  SerProThrAlaArgLeuThrAlaAlaAlaLeuAlaAlaLeuPheTyrCysAlaValAla  59
              :::|||  ::
Db 18043  TCCAGACGGCCGGCGTCTGGTTGCTATTTTTATTTATGGGGTTCTGGTTATTGCTGTCATT  18102

Qy    60  TrpAlaGlyLeuGlyArgIleAspIleValAlaSerAlaSerArgLysIleValProGly  79
              |||  ||||:|:|||||||||||||
Db 18103  TTATCTGTTTTAGGTCAGGTGAAATGTTGCCACTGCAAATGGGAAATTAACACTAAGT  18162

Qy    80  AspArgValLysLeuValGlnProLeuGluValAlaGlyValValArgAlaThrHisValArg  99
              |||      ::|||   :::|||  |:::||||||
Db 18163  GGGCGCAGCAAAGAAATTAAACCTATTGAAAACTCAATAGTTAAAGAAATTATCGTAAAA  18222

Qy   100  AspGlyGlnThrValLysAlaGlyGluIleLeuIleGluLeuAspProPheAlaGlyGly  119
              :::|||  ::::||  |||||||
Db 18223  GAAGGAGAGTCAGTCCGAAAGGGGATGTGTTATTAAAGCTTACAGCACTGGGAGCTGAA  18282

Qy   :20  ValAsp 121
              |||
Db 18283  GCTGAT 18288
```

RESULT 3
US-09-956-004-79/C
; Sequence 79, Application US/09956004
; Patent No. US2002007259SA1
GENERAL INFORMATION:
, APPLICANT: Patrick J. Dillon et al.
TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands
CURRENT APPLICATION NUMBER: US/09/956,004
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: 08/976,259
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/061,953
PRIOR FILING DATE: 1997-10-14
PRIOR APPLICATION NUMBER: 60/031,626
PRIOR FILING DATE: 1996-11-22

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 11:14:20 ; Search time 57 Seconds
(without alignments)
960.202 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 608
Sequence: 1 LSSLSPRMKSAREVVAVGGK.........KAGEILIELDPFAGGVDVAT 124

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame+_p2n.model -DEV=xlh
Q=/cgn2_1/USPTO_spool/US09484577/runat_23102003_111411_18018/app_query.fasta_1.263
DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=C
LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human4c.cdi
LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09484577 @CGN_1.1.56 @runat_23102003_111411_18018 -NCPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

	Result No.	Score	Query Match	Length	DB	ID	Description
c	1	215	35.4	28804	2	US-08-592-874-1	Sequence 1, Appli
c	2	215	35.4	28804	3	US-09-096-942-2	Sequence 2, Appli
c	3	215	35.4	28804	3	US-09-096-867-2	Sequence 2, Appli
	4	169	27.8	5120	3	US-08-772-270A-6	Sequence 6, Appli
	5	169	27.8	8370	2	US-08-488-706-1	Sequence 1, Appli
	6	169	27.8	8370	4	US-09-062-126-1	Sequence 1, Appli
	7	168	27.6	7721	3	US-08-772-270A-14	Sequence 14, Appl
c	8	154	25.3	3576	4	US-08-976-259-79	Sequence 79, Appl
	9	117	19.2	1236	4	US-09-328-352-2593	Sequence 2593, Ap
	10	109	17.9	1263	4	US-09-252-991A-13467	Sequence 13467, A
	11	109	17.9	1263	4	US-09-252-991A-13658	Sequence 13658, A
	12	104.5	17.2	1380	4	US-09-252-991A-3550	Sequence 3550, Ap
	13	104.5	17.2	1947	4	US-09-252-991A-3557	Sequence 3557, Ap
c	14	104	17.1	852	4	US-09-252-991A-11544	Sequence 11544, A
c	15	104	17.1	1608	4	US-09-252-991A-11665	Sequence 11665, A
	16	104	17.1	1608	4	US-09-252-991A-11769	Sequence 11769, A
	17	103	16.9	25165	5	US-09-453-702B-39	Sequence 39, Appli
	18	99	16.3	4465	2	US-08-620-605D-1	Sequence 1, Appli
	19	99	16.3	4547	2	US-09-005-232A-1	Sequence 1, Appli
	20	90	14.8	1302	4	US-09-491-772-5	Sequence 5, Appli
	21	90	14.8	8580	4	US-09-491-772-1	Sequence 1, Appli
c	22	87.5	14.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	23	87.5	14.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	24	82	13.5	885	4	US-09-252-991A-11189	Sequence 11189, A
c	25	82	13.5	1245	4	US-09-252-991A-11257	Sequence 11257, A
	26	82	13.5	1386	4	US-09-252-991A-10994	Sequence 10994, A
	27	80	13.2	1011	4	US-09-328-352-4076	Sequence 4076, Ap
	28	79.5	13.1	945	4	US-09-252-991A-3765	Sequence 3765, Ap
c	29	79.5	13.1	1383	4	US-09-252-991A-3609	Sequence 3609, Ap
	30	79.5	13.1	1533	4	US-09-252-991A-8687	Sequence 8687, Ap
c	31	78	12.8	2398	4	US-09-453-702B-260	Sequence 260, Appl
	32	77	12.7	49272	1	US-08-614-770A-1	Sequence 1, Appli
	33	76	12.5	1251	4	US-09-252-991A-444	Sequence 444, App
	34	76	12.5	1362	4	US-09-252-991A-417	Sequence 417, App
	35	76	12.5	2256	4	US-09-252-991A-16502	Sequence 16502, A
c	36	75.5	12.4	1321	4	US-09-175-658B-20	Sequence 20, Appl
c	37	75.5	12.4	1543	4	US-09-364-230-17	Sequence 17, Appl
	38	74.5	12.3	3603	1	US-08-188-582-15	Sequence 15, Appl
	39	74.5	12.3	3603	1	US-08-646-715-15	Sequence 15, Appl
	40	73	12.0	399	4	US-09-252-991A-14158	Sequence 14158, A
c	41	73	12.0	1161	4	US-09-252-991A-14117	Sequence 14117, A
c	42	73	12.0	1233	4	US-09-252-991A-14019	Sequence 14019, A
	43	73	12.0	1239	4	US-09-252-991A-8514	Sequence 8514, Ap
	44	73	12.0	1788	4	US-09-252-991A-10460	Sequence 10460, A
c	45	73	12.0	1911	4	US-09-252-991A-8331	Sequence 8331, Ap

ALIGNMENTS

RESULT 1
US-08-592-874-1/C
; Sequence 1, Application US/08592874
; Patent No. 5854034
GENERAL INFORMATION:
APPLICANT: POLLOCK, THOMAS J.
APPLICANT: YAMAZAKI, MOTOHIDE
APPLICANT: THORNE, LINDA
APPLICANT: MIKOLAJCZAK, MARCIA
APPLICANT: ARMENTROUT, RICHARD W.
TITLE OF INVENTION: DNA SEGMENTS AND METHODS FOR INCREASING
TITLE OF INVENTION: POLYSACCHARIDE PRODUCTION
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: JULES E. GOLDBERG
STRET: 261 MADISON AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016-2391
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,874
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/377,440
FILING DATE: 24-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: GOLDBERG, JULES E.
REGISTRATION NUMBER: 24,408
TELECOMMUNICATION INFORMATION:

;   TELEPHONE: 212-986-4090
;   TELEFAX: 212-818-9479
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 28804 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: unknown
;   TOPOLOGY: unknown
;   MOLECULE TYPE: DNA (genomic)
;   FRAGMENT TYPE: N-terminal
US-08-592-874-1

Alignment Scores:
Pred. No.:   1.35e-15   Length: 28804
Score: 215.00   Matches: 44
Percent Similarity: 60.82%   Conservative: 15
Best Local Similarity: 45.36%   Mismatches: 38
Query Match: 35.36%   Indels: 0
                                      Gaps: 0

US-09-484-577A-4 (1-124) x US-08-592-874-1 (1-28804)

Qy   27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAlaArgLeuThr 46
          ||||||||||||||||||| |||::||||
Db 18808 TTCCTGCCGGCCGCGCTCGAGATCATCGAACGTCCGGTGTCCGGCCGCGCCTCACC 18749

Qy   47 AlaAlaLeuLeuAlaAlaLeuPheTyrCysAlaValAlaLeuTrpAlaGlyLeuGlyArgIle 66
          ::::: ||
Db 18748 GGCCGGGTGATGTGGCGCGGACTGGCGATCACCACCGCCTGGCTGGCAATCGGCCGCGTG 18689

Qy   67 AspIleValAlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGln 86
          :::|| |||||| |||
Db 18688 GAAGTGGTGGCGCCGACGCAGGGCCGCATCGCCCGGCGAGACCAAGATCGTGCAG 18629

Qy   87 ProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
          ::::::|||||| || |||
Db 18628 TCCCCCGAAAGCGGTATCGTCCGGCGCATTCTGGTGGGCGAGAAGGTTCAGAAG 18569

Qy   107 GlyGluIleLeuIleGluLeuAspProPheAlaGlyGlyValValAspValAla 123
          ||||::::|||||| :: |||
Db 18568 GGCCAGGTGCTGATCACGCTCGATCCCACCATGTCGGCAGCCGAAGCCGCG 18518

RESULT 2
US-09-096-942-2/c
; Sequence 2, Application US/09096942
; Patent No. 6027925
; GENERAL INFORMATION:
;   APPLICANT: Pollock, Thomas J
;   APPLICANT: Mikolajczak, Marcia
;   APPLICANT: Yamazaki, Motohide
;   APPLICANT: Thorne, Linda
;   APPLICANT: Armentrout, Richard W
;   TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
;   TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
;   FILE REFERENCE: seq list for appl filed from pro. appl
;   CURRENT APPLICATION NUMBER: US/09/096,942
;   CURRENT FILING DATE: 1998-06-12
;   EARLIER APPLICATION NUMBER: 60/049,428
;   EARLIER FILING DATE: 1997-06-12
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 28804
;   TYPE: DNA
;   ORGANISM: Sphingomonas sp. S88
US-09-096-942-2

Alignment Scores:
Pred. No.:   1.35e-15   Length: 28804
Score: 215.00   Matches: 44
Percent Similarity: 60.82%   Conservative: 15
Best Local Similarity: 45.36%   Mismatches: 38
Query Match: 35.36%   Indels: 0

DB: 3   Gaps: 0

US-09-484-577A-4 (1-124) x US-09-096-942-2 (1-28804)

Qy   27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAlaArgLeuThr 46
          ||||||||||||||||||| |||::||||
Db 18808 TTCCTGCCGGCCGCGCTCGAGATCATCGAACGTCCGGTGTCCGGCCGCGCCTCACC 18749

Qy   47 AlaAlaLeuLeuAlaAlaLeuPheTyrCysAlaValAlaLeuTrpAlaGlyLeuGlyArgIle 66
          ::::: ||
Db 18748 GGCCGGGTGATGTGGCGCGGACTGGCGATCACCACCGCCTGGCTGGCAATCGGCCGCGTG 18689

Qy   67 AspIleValAlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGln 86
          :::|| |||||| |||
Db 18688 GAAGTGGTGGCGCCGACGCAGGGCCGCATCGCCCGGCGAGACCAAGATCGTGCAG 18629

Qy   87 ProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
          ::::::|||||| || |||
Db 18628 TCCCCCGAAAGCGGTATCGTCCGGCGCATTCTGGTGGGCGAGAAGGTTCAGAAG 18569

Qy   107 GlyGluIleLeuIleGluLeuAspProPheAlaGlyGlyValValAspValAla 123
          ||||::::|||||| :: |||
Db 18568 GGCCAGGTGCTGATCACGCTCGATCCCACCATGTCGGCAGCCGAAGCCGCG 18518

RESULT 3
US-09-096-867-2/c
; Sequence 2, Application US/09098867
; Patent No. 6030817
; GENERAL INFORMATION:
;   APPLICANT: Pollock, Thomas J
;   APPLICANT: Mikolajczak, Marcia
;   APPLICANT: Yamazaki, Motohide
;   APPLICANT: Thorne, Linda
;   APPLICANT: Armentrout, Richard W
;   TITLE OF INVENTION: Production of Xanthan Gum by Sphingomonas Bacteria
;   TITLE OF INVENTION: Carrying Genes from Xanthomonas Campestris
;   FILE REFERENCE: seq list for appl filed from pro. appl
;   CURRENT APPLICATION NUMBER: US/09/096,867
;   CURRENT FILING DATE: 1998-06-11
;   EARLIER APPLICATION NUMBER: 60/049,428
;   EARLIER FILING DATE: 1997-06-12
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 28804
;   TYPE: DNA
;   ORGANISM: Sphingomonas sp. S88
US-09-096-867-2

Alignment Scores:
Pred. No.:   1.35e-15   Length: 28804
Score: 215.00   Matches: 44
Percent Similarity: 60.82%   Conservative: 15
Best Local Similarity: 45.36%   Mismatches: 38
Query Match: 35.36%   Indels: 0
DB: Gaps: 0

US-09-484-577A-4 (1-124) x US-09-096-867-2 (1-28804)

Qy   27 PheLeuProAlaAlaLeuGluIleValGluThrProSerProThrAlaArgLeuThr 46
          ||||||||||||||||||| |||::||||
Db 18808 TTCCTGCCGGCCGCGCTCGAGATCATCGAACGTCCGGTGTCCGGCCGCGCCTCACC 18749

Qy   47 AlaAlaLeuLeuAlaAlaLeuPheTyrCysAlaValAlaLeuTrpAlaGlyLeuGlyArgIle 66
          ::::: ||
Db 18748 GGCCGGGTGATGTGGCGCGGACTGGCGATCACCACCGCCTGGCTGGCAATCGGCCGCGTG 18689

Qy   67 AspIleValAlaSerAlaSerArgLysIleValProGlyAspArgValLysLeuValGln 86
          :::|| |||||| |||
Db 18688 GAAGTGGTGGCGCCGACGCAGGGCCGCATCGCCCGGCGAGACCAAGATCGTGCAG 18629

Qy   87 ProLeuGluValGlyValValArgAlaThrHisValArgAspGlyGlnThrValLysAla 106
          ::::::|||||| || |||
Db 18628 TCCCCCGAAAGCGGTATCGTCCGGCGCATTCTGGTGGGCGAGAAGGTTCAGAAG 18569

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

Run on: October 23, 2003, 14:46:07 ; Search time 2169 Seconds
                                          (without alignments)
                                          9449.388 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actcccagcctctcaccga.........cagaatgcggcgatgatcat 501

Scoring table: OLIGO_NUC
             Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2016740

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :   GenEmbl:*
      1: gb_ba:*
      2: gb_htg:*
      3: gb_in:*
      4: gb_om:*
      5: gb_ov:*
      6: gb_pat:*
      7: gb_ph:*
      8: gb_pl:*
      9: gb_pr:*
     10: gb_ro:*
     11: gb_sts:*
     12: gb_sy:*
     13: gb_un:*
     14: gb_vi:*
     15: em_ba:*
     16: em_fun:*
     17: em_hum:*
     18: em_in:*
     19: em_mu:*
     20: em_om:*
     21: em_or:*
     22: em_ov:*
     23: em_pat:*
     24: em_ph:*
     25: em_pl:*
     26: em_ro:*
     27: em_sts:*
     28: em_un:*
     29: em_vi:*
     30: em_htg_hum:*
     31: em_htg_inv:*
     32: em_htg_other:*
     33: em_htg_mus:*
     34: em_htg_pln:*
     35: em_htg_rod:*
     36: em_htg_mam:*
     37: em_htg_vrt:*
     38: em_sy:*
     39: em_htgo_hum:*
     40: em_htgo_mus:*
     41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.		Score	% Query Match	Length	DB	ID	Description
1	c	24	4.8	60	6	AR231929	AR231929 Sequence
2	c	24	4.8	60	6	AR231929	AR231929 Sequence
3		24	4.8	60	6	AR231931	AR231931 Sequence
4	c	24	4.8	60	6	AR231931	AR231931 Sequence
5	c	24	4.8	60	6	AX464727	AX464727 Sequence
6	c	24	4.8	60	6	AX464727	AX464727 Sequence
7		24	4.8	60	6	AX464729	AX464729 Sequence
8	c	24	4.8	60	6	AX464729	AX464729 Sequence
9		21	4.2	24	6	A68884	A68884 Sequence 55
10		21	4.2	24	6	AR055096	AR055096 Sequence
11		21	4.2	24	6	AR055098	AR055098 Sequence
12		21	4.2	24	6	AR068437	AR068437 Sequence
13		21	4.2	24	6	AR068439	AR068439 Sequence
14		21	4.2	24	6	AR121280	AR121280 Sequence
15		21	4.2	24	6	AR231914	AR231914 Sequence
16		21	4.2	24	6	AX464709	AX464709 Sequence
17		21	4.2	24	6	BD062984	BD062984 DNA and s
18		21	4.2	24	6	I13333	I13333 Sequence 7
19		21	4.2	24	6	I19026	I19026 Sequence 7
20		21	4.2	28	6	AR035013	AR035013 Sequence
21		21	4.2	28	6	AR082097	AR082097 Sequence
22		21	4.2	28	6	AR118488	AR118488 Sequence
23		21	4.2	28	6	AR151056	AR151056 Sequence
24		21	4.2	28	6	AR218212	AR218212 Sequence
25		21	4.2	28	6	AR222652	AR222652 Sequence
26		21	4.2	28	6	AR231690	AR231690 Sequence
27		20	4.0	23	6	AX194017	AX194017 Sequence
28		20	4.0	24	6	A68887	A68887 Sequence 58
29		20	4.0	24	6	AR034959	AR034959 Sequence
30		20	4.0	24	6	AR034973	AR034973 Sequence
31		20	4.0	24	6	AR082043	AR082043 Sequence
32		20	4.0	24	6	AR082057	AR082057 Sequence
33		20	4.0	24	6	AR082102	AR082102 Sequence
34		20	4.0	24	6	AR082104	AR082104 Sequence
35		20	4.0	24	6	AR100939	AR100939 Sequence
36		20	4.0	24	6	AR100941	AR100941 Sequence
37		20	4.0	24	6	AR100943	AR100943 Sequence
38		20	4.0	24	6	AR100945	AR100945 Sequence
39		20	4.0	24	6	AR100965	AR100965 Sequence
40		20	4.0	24	6	AR100979	AR100979 Sequence
41		20	4.0	24	6	AR101019	AR101019 Sequence
42		20	4.0	24	6	AR101021	AR101021 Sequence
43		20	4.0	24	6	AR118434	AR118434 Sequence
44		20	4.0	24	6	AR118448	AR118448 Sequence
45		20	4.0	24	6	AR130747	AR130747 Sequence
46		20	4.0	24	6	AR151002	AR151002 Sequence
47		20	4.0	24	6	AR151016	AR151016 Sequence
48		20	4.0	24	6	AR199145	AR199145 Sequence
49		20	4.0	24	6	AR199149	AR199149 Sequence
50		20	4.0	24	6	AR199163	AR199163 Sequence
51		20	4.0	24	6	AR218158	AR218158 Sequence
52		20	4.0	24	6	AR218172	AR218172 Sequence
53		20	4.0	24	6	AR222598	AR222598 Sequence
54		20	4.0	24	6	AR222612	AR222612 Sequence
55		20	4.0	24	6	AR231636	AR231636 Sequence
56		20	4.0	24	6	AR231650	AR231650 Sequence
57		20	4.0	24	6	BD062987	BD062987 DNA and s
58		20	4.0	24	6	BD080944	BD080944 Detection
59		20	4.0	24	6	AR035015	AR035015 Sequence
60		20	4.0	48	6	AR082099	AR082099 Sequence
61		20	4.0	48	6	AR118490	AR118490 Sequence
62		20	4.0	48	6	AR151058	AR151058 Sequence
63		20	4.0	48	6	AR218214	AR218214 Sequence
64		20	4.0	48	6	AR222654	AR222654 Sequence
65		20	4.0	48	6	AR231692	AR231692 Sequence

```
 66  20  4.0  6  AR035017  Sequence
 67  20  4.0  6  AR082101  Sequence
 68  20  4.0  6  AR118492  Sequence
 69  20  4.0  6  AR151060  Sequence
 70  20  4.0  6  AR218216  Sequence
 71  20  4.0  6  AR222656  Sequence
 72  20  4.0  6  AR231694  Sequence
 73  19  3.8  6  AR121274  Sequence
 74  19  3.8  6  AR139761  Sequence
 75  19  3.8  6  AR165220  Sequence
 76  19  3.8  6  AR201391  Sequence
 77  19  3.8  6  AR208439  Sequence
 78  19  3.8  6  AR225774  Sequence
 79  19  3.8  6  AR230325  Sequence
 80  19  3.8  6  AR243315  Sequence
 81  19  3.8  6  AR264180  Sequence
 82  19  3.8  6  AR279223  Sequence
 83  19  3.8  6  AR310020  Sequence
 84  19  3.8  6  AX023539  Sequence
 85  19  3.8  6  AX301748  Sequence
 86  19  3.8  6  AX347891  Sequence
 87  19  3.8  6  AX348098  Sequence
 88  19  3.8  6  AX384667  Sequence
 89  19  3.8  6  AX459947  Sequence
 90  19  3.8  6  AX470041  Sequence
 91  19  3.8  6  BD016439  Gene deri
 92  19  3.8  6  BD080835  Asthma as
 93  19  3.8  6  BD082003  Asthma as
 94  19  3.8  6  BD086207  Transcrip
 95  19  3.8  6  BD106876  Isolated
 96  19  3.8  6  I13327    Sequence 1
 97  19  3.8  6  I19020    Sequence 1
 98  19  3.8  6  I27658    Sequence 13
 99  19  3.8  6  I27653    Sequence 8
100  19  3.8  6  I27649    Sequence 4
```

ALIGNMENTS

RESULT 1
AR231929
```
LOCUS       AR231929                  60 bp    DNA        linear   PAT 20-DEC-2002
DEFINITION  Sequence 19 from patent US 6455255.
ACCESSION   AR231929
VERSION     AR231929.1  GI:27273523
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Birkenmeyer,L.G., Leary,T.P., Muerhoff,A.S., Desai,S.M. and
            Mushahwar,I.K.
  TITLE     Method of performing subtractive hybridization using RDA
  JOURNAL   Patent: US 6455255-A 19 24-SEP-2002;
  FEATURES  Location/Qualifiers
     source          1..60
                     /organism="unknown"
BASE COUNT    11 a   17 c   17 g   11 t   4 others
ORIGIN

Query Match       4.8%;  Score 24;  DB 6;  Length 60;
Best Local Similarity  100.0%;  Pred. No. 0.3;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ACTCTCCAGCCTCTCACCGAGGAT 24
              ||||||||||||||||||||||||
Db          4 ACTCTCCAGCCTCTCACCGAGGAT 27
```

RESULT 2
AR231929/c
```
LOCUS       AR231929                  60 bp    DNA        linear   PAT 20-DEC-2002
DEFINITION  Sequence 19 from patent US 6455255.
ACCESSION   AR231929
VERSION     AR231929.1  GI:27273523
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Birkenmeyer,L.G., Leary,T.P., Muerhoff,A.S., Desai,S.M. and
            Mushahwar,I.K.
  TITLE     Method of performing subtractive hybridization using RDA
  JOURNAL   Patent: US 6455255-A 19 24-SEP-2002;
  FEATURES  Location/Qualifiers
     source          1..60
                     /organism="unknown"
BASE COUNT    11 a   17 c   17 g   11 t   4 others
ORIGIN

Query Match       4.8%;  Score 24;  DB 6;  Length 60;
Best Local Similarity  100.0%;  Pred. No. 0.3;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ACTCTCCAGCCTCTCACCGAGGAT 24
              ||||||||||||||||||||||||
Db         57 ACTCTCCAGCCTCTCACCGAGGAT 34
```

RESULT 3
AR231931
```
LOCUS       AR231931                  60 bp    DNA        linear   PAT 20-DEC-2002
DEFINITION  Sequence 21 from patent US 6455255.
ACCESSION   AR231931
VERSION     AR231931.1  GI:27273525
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Birkenmeyer,L.G., Leary,T.P., Muerhoff,A.S., Desai,S.M. and
            Mushahwar,I.K.
  TITLE     Method of performing subtractive hybridization using RDA
  JOURNAL   Patent: US 6455255-A 21 24-SEP-2002;
  FEATURES  Location/Qualifiers
     source          1..60
                     /organism="unknown"
BASE COUNT    11 a   17 c   17 g   11 t   4 others
ORIGIN

Query Match       4.8%;  Score 24;  DB 6;  Length 60;
Best Local Similarity  100.0%;  Pred. No. 0.3;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ACTCTCCAGCCTCTCACCGAGGAT 24
              ||||||||||||||||||||||||
Db          4 ACTCTCCAGCCTCTCACCGAGGAT 27
```

RESULT 4
AR231931/c
```
LOCUS       AR231931                  60 bp    DNA        linear   PAT 20-DEC-2002
DEFINITION  Sequence 21 from patent US 6455255.
ACCESSION   AR231931
VERSION     AR231931.1  GI:27273525
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Birkenmeyer,L.G., Leary,T.P., Muerhoff,A.S., Desai,S.M. and
            Mushahwar,I.K.
  TITLE     Method of performing subtractive hybridization using RDA
  JOURNAL   Patent: US 6455255-A 21 24-SEP-2002;
  FEATURES  Location/Qualifiers
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: October 23, 2003, 15:26:52 ; Search time 1679 Seconds
                                          (without alignments)
                                          7252.259 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 acctccagcctccaccga........cagaatgcggcgatgatcat 501

Scoring table: OLIGO_NUC
             Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4672214

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :   EST:*
             1: em_estba:*
             2: em_esthum:*
             3: em_estin:*
             4: em_estmu:*
             5: em_estov:*
             6: em_estpl:*
             7: em_estro:*
             8: em_hcc:*
             9: gb_est1:*
            10: gb_est2:*
            11: gb_htc:*
            12: gb_est3:*
            13: gb_est4:*
            14: gb_est5:*
            15: em_estfun:*
            16: em_estom:*
            17: em_gss_hum:*
            18: em_gss_inv:*
            19: em_gss_pln:*
            20: em_gss_vrt:*
            21: em_gss_fun:*
            22: em_gss_mam:*
            23: em_gss_mus:*
            24: em_gss_pro:*
            25: em_gss_rod:*
            26: em_gss_phg:*
            27: em_gss_vrl:*
            28: gb_gss1:*
            29: gb_gss2:*

No.	Flag	Score	%	Length	DB	ID	Description
5	c	19	3.8	100	10	BG599358	BG599358 ESTO0004
6		19	3.8	109	10	BG315516	BG315516 PO3.0.187
7		19	3.8	121	10	BG315546	BG315546 PO3.0.293
8		19	3.8	131	10	BG315530	BG315530 PO3.0.237
9		19	3.8	132	10	BE644247	BE644247 STRF004 U
10		19	3.8	189	9	AF027371	AF027371 AF027371
11		19	3.8	214	9	AF027372	AF027372 AF027372
12		19	3.8	219	9	AF027370	AF027370 AF027370
13		19	3.8	229	9	AF027374	AF027374 AF027374
14		19	3.8	241	9	AJ414389	AJ414389 AJ414389
15	c	19	3.8	243	9	AJ414388	AJ414388 AJ414388
16	c	18	3.6	138	9	AF091520	AF091520 AF091520
17		18	3.6	224	12	BG946550	BG946550 PM3-Kf004
18		17	3.4	103	9	AA234634	AA234634 zr78d10.r
19		17	3.4	107	28	BH216229	BH216229 1006041C0
20		17	3.4	127	28	BE588545	BE588545 193860 BA
21		17	3.4	163	12	AI192272	AI192272 qc97c08.x
22		17	3.4	167	12	BM706517	BM706517 UI-E-CR1-
23		17	3.4	167	14	T59467	T59467 yb63d03.r1
24		17	3.4	194	12	BM719924	BM719924 UI-B-EJO-
25		17	3.4	216	12	N90915	N90915 zb11e03.s1
26		17	3.4	227	9	AI910087	AI910087 RC-BT236-
27	c	17	3.4	232	14	CB393844	CB393844 OSTR127G2
28		17	3.4	235	14	AA841745	AA841745 MB3D6V6F1
29		17	3.4	236	14	CB389983	CB389983 OSTF127G2
30		17	3.4	240	10	AI561896	AI561896 vj49a08.x
31		17	3.4	244	14	BG360850	BG360850 947042G10
32	c	17	3.4	247	14	CB031588	CB031588 TgESTzyc0
33		16	3.2	87	28	AZ594197	AZ594197 1M0406I14
34		16	3.2	104	12	BI396353	BI396353 94047D12
35	c	16	3.2	110	10	BG354695	BG354695 ESTO5 Hum
36		16	3.2	114	10	BW300094	BW300094 BW300094
37	c	16	3.2	117	13	AZ893616	AZ893616 RPCI-24-1
38	c	16	3.2	119	28	AL479835	AL479835 T. brucei
39	c	16	3.2	135	12	AA463183	AA463183 vf93a04.r
40		16	3.2	141	12	BJ234310	BJ234310 BJ234310
41		16	3.2	144	10	BE497583	BE497583 WHE0754_F
42		16	3.2	149	29	CC178168	CC178168 XC354_Bay
43		16	3.2	159	29	CB066573	CB066573 EST0008 G
44		16	3.2	160	10	BE497606	BE497606 WHE0754_D
45		16	3.2	166	13	BY008185	BY008185 BY008185
46		16	3.2	168	13	BI711777	BI711777 id97e05.Y
47		16	3.2	169	9	AV414210	AV414210 AV414210
48	c	16	3.2	171	12	AL121138	AL121138 DKFZp762I
49		16	3.2	179	12	BM174574	BM174574 Tm_ad_31B
50	c	16	3.2	179	29	BX003993	BX003993 ArabIdops
51		16	3.2	180	10	BE438922	BE438922 BCD102_W
52		16	3.2	181	28	CA347516	CA347516 678657_NC
53	c	16	3.2	189	28	BH863430	BH863430 SALK_0938
54	c	16	3.2	192	14	CB473406	CB473406 sn69_F09.
55		16	3.2	194	14	AF027371	AF027371 AF027371
56	c	16	3.2	196	28	AZ922877	AZ922877 SLCot5G07
57	c	16	3.2	196	9	AI903647	AI903647 QV-BT032-
58		16	3.2	198	14	AA066211	AA066211 mm43g04.r
59		16	3.2	198	12	AA931123	AA931123 om87f11.s
60		16	3.2	202	9	BQ401278	BQ401278 I091043E0
61		16	3.2	206	13	CA698323	CA698323 wlk8.pk00
62	c	16	3.2	214	14	CA659563	CA659563 wlm1.pk00
63		16	3.2	221	14	AL002792	FR0011531 F.rubripe
64	c	16	3.2	226	12	BJ322497	BJ322497 BJ322497
65	c	16	3.2	232	22	W83952	W83952 mf34f08.r1
66	c	16	3.2	238	14	AV416808	AV416808 AV416808
67	c	16	3.2	232	9	BI042670	BI042670 MR3-OT018
68		16	3.2	239	14	CA958302	CA958302 px58b04.Y
69		16	3.2	243	14	BX257343	BX257343 BX257343
70		15	3.0	60	13	BB148038	BB148038 BB148038
71		15	3.0	107	13	AL754758	AL754758 Arabidops
72	c	15	3.0	116	29	BI683271	BI683271 464597 MA
73		15	3.0	116	12	AW361611	AW361611 MR2-CT025
74		15	3.0	121	13	BU998747	BU998747 HI12A18r
75	c	15	3.0	123	14	CB929267	CB929267 ABA1_41_D
76	c	15	3.0	125	14	BJ211548	BJ211548 BJ211548
77	c	15	3.0	125	28	BH192273	BH192273 TC3-71J24

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.		Score	% Query Match	Length	DB	ID	Description
C	1	24	4.8	131	9	AF091519	AF091519 AF091519
C	2	22	4.4	104	14	CA453170	CA453170 EST01 Hum
C	3	20	4.0	178	28	AZ924992	AZ924992 Bt44 Homo
C	4	19	3.8	82	10	BG599357	BG599357 EST00003

```
c   78    15    3.0    BF171536    127    10    BF171536   PCL2892 M
c   79    15    3.0    BH195962    127    28    BH195962   TC3-8F12.
c   80    15    3.0    BE183236    129    28    BE183236   CX2-HT065
c   81    15    3.0    BJ303021    129    12    BJ303021   BJ303021
c   82    15    3.0    AW556423    130     9    AW556423   L0268B02-
c   83    15    3.0    CA533574    135    14    CA533574   C0403A10-
c   84    15    3.0    BE238693    139    10    BE238693   946007A04
c   85    15    3.0    BF484273    140    10    BF484273   WHE2321 C
c   86    15    3.0    BH230411    144    28    BH230411   1006157G1
c   87    15    3.0    AW176310    146     9    AW176310   QV0-BT022
c   88    15    3.0    AZ050374    147    28    AZ050374   GSSTc1:21
c   89    15    3.0    BH194741    148    28    BH194741   TC3-70A23
c   90    15    3.0    AZ802087    150    28    AZ802087   2X0060A19
c   91    15    3.0    AW854662    152    10    AW854662   MRI-CT025
c   92    15    3.0    BM823569    157    12    BM823569   K-EST0094
c   93    15    3.0    AA573678    160     9    AA573678   nf42h02.s
c   94    15    3.0    AL504155    165     9    AL504155   AL504155
c   95    15    3.0    AU209703    167    28    AU209703   AU209703
c   96    15    3.0    BE995856    167    10    BE995856   UI-M-CGOp
c   97    15    3.0    BW263975    168    13    BW263975   BW263975
c   98    15    3.0    AZ115539    168    28    AZ115539   RPCI-23-4
c   99    15    3.0    AZ577177    168    28    AZ577177   05g07 Sho
c  100    15    3.0    T12301      170    14    T12301     A038R Heart
```

ALIGNMENTS

```
RESULT 1
AF091519
LOCUS       AF091519               131 bp    mRNA    linear   EST 10-NOV-1996
DEFINITION  AF091519 Human nasopharyngeal carcinoma cell line HNE1 Homo sapiens
            cDNA clone LZ17, mRNA sequence.
ACCESSION   AF091519
VERSION     AF091519.1  GI:3851150
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 131)
AUTHORS     Zhan,F.H. and Li,G.Y.
TITLE       Differentially expressed cDNA sequence
JOURNAL     Unpublished
COMMENT     Contact: Zhan FH
            Cancer Research Institute
            88 Xiangya Road, Changsha, Hunan, 410078, China.
FEATURES             Location/Qualifiers
     source          1..131
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="LZ17"
                     /tissue_type="nasopharyngeal carcinoma"
                     /cell_line="HNE1"
                     /clone_lib="Human nasopharyngeal carcinoma cell line HNE1"
BASE COUNT    43 a     35 c     27 g     26 t
ORIGIN
```

```
Query Match             4.8%;  Score 24;  DB 9;  Length 131;
Best Local Similarity 100.0%;  Pred. No. 0.074;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ACTCTCCAGCCTCTCACCGAGGAT 24
          ||||||||||||||||||||||||
Db      2 ACTCTCCAGCCTCTCACCGAGGAT 25
```

```
RESULT 2
CA453170/c
LOCUS       CA453170               104 bp    mRNA    linear   EST 12-NOV-2002
DEFINITION  EST01 Human transformed 16HBE cDNA Homo sapiens cDNA clone CHL1,
            mRNA sequence.
```

```
ACCESSION   CA453170
VERSION     CA453170.1  GI:24934968
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 104)
AUTHORS     Lijin,Z.H. and Yiguo,J.
TITLE       Cloning of genes related to Chlorophyllin Antitransforming Against
            Anti-BPDE
JOURNAL     Zhongliu Fangzhi Zazhi 10 (1), 31-34 (2003)
COMMENT     Contact: Lijin Z
            Institute for Chemical Carcinogenesis
            GuangZhou Medical College
            195 Dongfengxi Road, Guangzhou 510182, China
            Tel: 8620 81340186
            Email: lijin_zhu@hotmail.com.
FEATURES             Location/Qualifiers
     source          1..104
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CHL1"
                     /cell_type="Bronchial epithelial cell"
                     /cell_line="16HBE"
                     /clone_lib="Human transformed 16HBE cDNA"
                     /note="Vector: pGEM-T; cDNA sequence from Homo sapiens
                     16HBE cells transformed by BPDE"
BASE COUNT    17 a     21 c     46 g     20 t
ORIGIN
```

```
Query Match             4.4%;  Score 22;  DB 14;  Length 104;
Best Local Similarity 100.0%;  Pred. No. 0.82;
Matches   22;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 ACTCTCCAGCCTCTCACCGAGG 22
          ||||||||||||||||||||||
Db     76 ACTCTCCAGCCTCTCACCGAGG 55
```

```
RESULT 3
AZ924992/c
LOCUS       AZ924992               178 bp    DNA    linear   GSS 22-MAR-2001
DEFINITION  Bt44 Homo sapiens genomic DNA random plasmid subclone for CpG
            island Homo sapiens genomic clone Bt44 5', genomic survey sequence.
ACCESSION   AZ924992
VERSION     AZ924992.1  GI:13433275
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 178)
AUTHORS     Warachit,J., Porntanakasem,W. and Mutirangura,A.
TITLE       Identification of differentially methylated sequence between white
            blood cells and sperm
JOURNAL     Unpublished
COMMENT     Contact: Warachit J.
            Medical Oncology Unit, Faculty of Medicine
            Chulalongkorn University
            Prathumwan Rd., Bangkok, Thailand, 10330
            Tel: 662-256-4532
            Fax: 662-256-4534
            Email: jiranan_w@hotmail.com
            Plate: 4
            Seq primer: T7
            Class: random plasmid subclone
            High quality sequence stop: 178.
FEATURES             Location/Qualifiers
     source          1..178
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: October 23, 2003, 14:45:07 ; Search time 233 Seconds
                                (without alignments)
                                5804.371 Million cell updates/sec

Title: US-09-484-577A-3
Perfect score: 501
Sequence: 1 actctccagcctctcaccga.........cagaatgcgcgatgatcat 501

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size :   0

Total number of hits satisfying chosen parameters: 3107778

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :   N_Geneseq_19Jun03:*
              1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
              2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
              3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
              4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
              5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
              6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
              7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
              8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
              9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
            10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
            11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
            12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
            13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
            14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
            15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
            16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
            17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
            18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
            19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
            20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
            21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
            22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
            23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
            24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
            25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	DB ID		Description
1	24	4.8	60	24	AAD32135		Sau3AI DNA fragmen
c 2	24	4.8	60	24	AAD32135		Sau3AI DNA fragmen
3	24	4.8	60	24	AAD32137		Sau3AI DNA fragmen
c 4	24	4.8	60	24	AAD32137		Sau3AI DNA fragmen
5	21	4.2	24	15	AAQ65644		Primer/adaptor uae
6	21	4.2	24	17	AAT13201		Primer R Bam 24 us
7	21	4.2	24	19	AAV03563		Probe RBam24 for s
8	21	4.2	24	20	AAZ24707		Adaptor sequence r

9	24	21	4.2	20	AAX21386	Primer R-Bam-24 fo
10	24	21	4.2	21	AAZ92275	Oligonucleotide #7
11	24	21	4.2	21	ABA88731	Oligonucleotide SE
12	24	22	4.2	21	ABA88733	Oligonucleotide SE
13	24	22	4.2	21	AAF24051	RDA primer R Bam24
14	24	22	4.2	21	AAI72946	Primer R Bam24, se
15	24	24	4.2	21	AAD32117	Sau3AI DNA fragmen
16	23	22	4.0	20	AAH23398	Nucleotide sequenc
17	24	22	4.0	20	AAT92397	Adapter primer RA2
18	24	18	4.0	20	AAT92398	Adapter primer RC2
19	24	18	4.0	20	AAT92401	Adapter primer RA2
20	24	18	4.0	20	AAT92403	Adapter primer RC2
21	24	19	4.0	20	AAV03566	Probe RMsp24 for s
22	24	19	4.0	20	AAV05432	Primer used in pro
23	24	19	4.0	20	AAV05434	Primer used in pro
24	24	19	4.0	20	AAV05436	Primer used in pro
25	24	19	4.0	20	AAV05438	Primer used in pro
26	24	19	4.0	20	AAV05440	Primer used in pro
27	24	19	4.0	20	AAV05400	Primer RA24 used i
28	24	19	4.0	20	AAV05401	Primer RC24 used i
29	24	19	4.0	20	AAV05404	Primer RA24-U used
30	24	19	4.0	20	AAV05406	Primer RC24-U used
31	24	20	4.0	20	AAZ24719	Adaptor sequence r
32	24	20	4.0	20	AAX24299	Human c1R BspHI-Ec
33	24	21	4.0	20	AAA72895	Yeast PCR primer #
34	24	21	4.0	20	AAA72897	Yeast PCR primer #
35	24	21	4.0	20	AAA72899	Yeast PCR primer #
36	24	21	4.0	20	AAA73386	Yeast PCR primer #
37	24	21	4.0	20	AAA73406	Linker RA23. Sacc
38	24	21	4.0	20	AAA73420	Yeast PCR primer R
39	24	21	4.0	20	AAA73460	Yeast PCR primer R
40	24	21	4.0	20	AAA73462	Yeast PCR primer R
41	24	21	4.0	20	AAA55868	Sequencing and PCR
42	24	21	4.0	20	AAA55870	Sequencing and PCR
43	24	21	4.0	20	AAA55872	Tsp 509 I adapter
44	24	21	4.0	20	AAA55874	Hind III adapter p
45	24	21	4.0	20	AAA55894	Adapter primer nuc
46	24	21	4.0	20	AAA55908	Adapter primer nuc
47	24	21	4.0	20	AAA55948	Adapter primer nuc
48	24	21	4.0	20	AAA55950	Adapter primer nuc
49	24	21	4.0	20	AAD00452	Oligonucleotide RX
50	24	21	4.0	20	AAZ38891	Sample adapter pri
51	24	21	4.0	20	AAZ38892	Sample adapter pri
52	24	22	4.0	20	AAF27247	RTaqI adapter, str
53	24	24	4.0	20	AAD45471	Sau3A I enzyme amp
54	24	24	4.0	20	AAD45473	BsaW I enzyme ampl
55	24	24	4.0	20	AAD45475	Tsp 509 I restrict
56	24	24	4.0	20	AAD45477	Hind III restricti
57	24	24	4.0	20	AAD45497	RA24 PCR primer us
58	24	24	4.0	20	AAD45511	RC24 PCR primer us
59	24	24	4.0	20	AAD45551	RA24-U PCR primer
60	24	24	4.0	20	AAD45553	RC24-U PCR primer
61	24	20	4.0	20	AAT92394	Antisense RNA prob
62	48	18	4.0	20	AAT92396	Antisense RNA prob
63	53	25	3.8	19	ABA00925	Non-biotinylated p
64	23	15	3.8	19	AAQ65638	Primer/adaptor use
65	24	16	3.8	19	AAT00032	HGBV DNA PCR prime
66	24	17	3.8	19	AAT39259	Primer, Lu3, used
67	24	17	3.8	19	AAT13195	Primer R Bgl 24 us
68	24	19	3.8	19	AAV64575	Human Fanconi anae
69	24	19	3.8	19	AAV27826	Representational d
70	24	19	3.8	19	AAV26812	PCR suppression pr
71	24	20	3.8	19	AAZ24701	Adaptor sequence r
72	24	20	3.8	19	AAZ24713	Adaptor sequence r
73	24	20	3.8	19	AAX58761	Oligonucleotide RB
74	24	20	3.8	19	AAX37656	Representational d
75	24	20	3.8	19	AAX21380	Primer R-Bgl-24 fo
76	24	20	3.8	19	AAA28819	Oligo for subtract
77	24	21	3.8	19	AAA55270	R Bgl 24 PCR prime
78	24	21	3.8	19	AAZ95366	RDA oligonucleotid
79	24	21	3.8	19	AAZ88523	Murine Ha3 gene pr
80	24	21	3.8	19	AAZ36285	Oligonucleotide pr
81	24	21	3.8	19	AAZ36629	PCR primer R-Bgl-2

```
82   19   3.8   AAI68204   24   22   Human osteoclast e
83   19   3.8   AAS14852   24   22   Mouse T cell deriv
84   19   3.8   AAD06783   24   22   R-Bgl-24 adaptor u
85   19   3.8   AAF24045   24   24   RDA primer R Bgl10
86   19   3.8   ABX04209   24   24   Representational d
87   19   3.8   AAL41512   24   24   Oligonucleotide in
88   19   3.8   AAL47921   24   24   Human chordrocyte
89   19   3.8   AAI72940   24   24   Primer R Bgl124, s
90   19   3.8   ABK47764   24   24   Human inflammation
91   19   3.8   AAD29765   24   24   PCR primer used in
92   19   3.8   AAD30621   24   24   PCR primer #1 used
93   19   3.8   AAD27127   24   24   Oligo #1 to clone
94   19   3.8   AAD23994   24   24   Adaptor #1 used in
95   19   3.8   ABA92366   24   24   Oligonucleotide ad
96   19   3.8   AAS17926   24   24   Representational d
97   19   3.8   AAS17613   24   24   Mouse G protein-co
98   19   3.8   ABX93698   24   25   R-Bgl adaptor olig
99   19   3.8   ABA00923   41   25   Biotinylated probe
100  19   3.8   AAT39254   45   17   PCR suppression ad
```

ALIGNMENTS

RESULT 1
AAD32135
ID   AAD32135 standard; DNA; 60 BP.
XX
AC   AAD32135;
XX
DT   18-JUN-2002 (first entry)
XX
DE   Sau3AI DNA fragment amplifying R-Bam 24 PCR primer #2.
XX
KW   Subtractive hybridisation; nucleic acid isolation technique; Sau3AI;
KW   visual identification; PCR; primer; ss.
XX
OS   Unidentified.
XX
PN   WO200210458-A2.
XX
PD   07-FEB-2002.
XX
PF   02-AUG-2001; 2001WO-US24480.
XX
PR   02-AUG-2000; 2000US-0631349.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Birkenmeyer LG, Leary TP, Muerhoff AS, Desai SM, Mushahwar IK;
XX
DR   WPI; 2002-269020/31.
XX
PT   Improved method for performing subtractive hybridization useful in
PT   nucleic acid isolation techniques, by employing Selective Primed
PT   Adaptive Driver-RDA, which utilizes a tester sample and a driver sample
PT   -
XX
PS   Example 2; Fig 1a; 67pp; English.
XX
CC   The invention relates to an improved method for performing subtractive
CC   hybridisation. The method involves using a tester sample and a driver
CC   sample to determine the presence of a nucleic acid sequence difference
CC   in the tester sample. The method is useful for performing subtractive
CC   hybridisation particularly for improving nucleic acid isolation
CC   techniques. The method may also be used for the visual identification of
CC   unique tester sequences. The present sequence is a PCR primer used for
CC   amplifying Sau3AI DNA fragment used in the exemplification of the
CC   invention.
XX
SQ   Sequence 60 BP; 11 A; 17 C; 17 G; 11 T; 4 other;

Query Match   4.8%; Score 24; DB 24; Length 60;

Best Local Similarity   100.0%; Pred. No. 0.036;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 ACTCTCCAGCCTCTCACCGAGGAT 24
        ||||||||||||||||||||||||
Db    4 ACTCTCCAGCCTCTCACCGAGGAT 27
```

RESULT 2
AAD32135/c
ID   AAD32135 standard; DNA; 60 BP.
XX
AC   AAD32135;
XX
DT   18-JUN-2002 (first entry)
XX
DE   Sau3AI DNA fragment amplifying R-Bam 24 PCR primer #2.
XX
KW   Subtractive hybridisation; nucleic acid isolation technique; Sau3AI;
KW   visual identification; PCR; primer; ss.
XX
OS   Unidentified.
XX
PN   WO200210458-A2.
XX
PD   07-FEB-2002.
XX
PF   02-AUG-2001; 2001WO-US24480.
XX
PR   02-AUG-2000; 2000US-0631349.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Birkenmeyer LG, Leary TP, Muerhoff AS, Desai SM, Mushahwar IK;
XX
DR   WPI; 2002-269020/31.
XX
PT   Improved method for performing subtractive
PT   nucleic acid isolation techniques, by employing Selective Primed
PT   Adaptive Driver-RDA, which utilizes a tester sample and a driver sample
PT   -
XX
PS   Example 2; Fig 1a; 67pp; English.
XX
CC   The invention relates to an improved method for performing subtractive
CC   hybridisation. The method involves using a tester sample and a driver
CC   sample to determine the presence of a nucleic acid sequence difference
CC   in the tester sample. The method is useful for performing subtractive
CC   hybridisation particularly for improving nucleic acid isolation
CC   techniques. The method may also be used for the visual identification of
CC   unique tester sequences. The present sequence is a PCR primer used for
CC   amplifying Sau3AI DNA fragment used in the exemplification of the
CC   invention.
XX
SQ   Sequence 60 BP; 11 A; 17 C; 17 G; 11 T; 4 other;

Query Match   4.8%; Score 24; DB 24; Length 60;
Best Local Similarity   100.0%; Pred. No. 0.036;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1 ACTCTCCAGCCTCTCACCGAGGAT 24
        ||||||||||||||||||||||||
Db   57 ACTCTCCAGCCTCTCACCGAGGAT 34
```

RESULT 3
AAD32137
ID   AAD32137 standard; DNA; 60 BP.
XX
AC   AAD32137;
XX
DT   18-JUN-2002 (first entry)
XX

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     October 23, 2003, 15:30:02 ; Search time 64 seconds
                                         (without alignments)
                                 3455.202 Million cell updates/sec

Title:          US-09-484-577A-3
Perfect score:  501
Sequence:       1 actctccagcctctcaccga........cagaatgcggcgatgatcat 501

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       569978 seqs, 220691566 residues

Word size :     0

Total number of hits satisfying chosen parameters:      874954

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :    Issued_Patents NA:*
              1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
              6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	24	4.8	60	4	US-09-631-349A-19	Sequence 19, Appl
2	24	4.8	60	4	US-09-631-349A-19	Sequence 19, Appl
3	24	4.8	60	4	US-09-631-349A-21	Sequence 21, Appl
4	24	4.8	60	4	US-09-631-349A-21	Sequence 21, Appl
5	21	4.2	24	1	US-07-974-447-7	Sequence 7, Appli
6	21	4.2	24	1	US-08-149-199-7	Sequence 7, Appli
7	21	4.2	24	2	US-08-858-767-1	Sequence 1, Appli
8	21	4.2	24	2	US-08-858-767-3	Sequence 3, Appli
9	21	4.2	24	2	US-08-863-028-1	Sequence 1, Appli
10	21	4.2	24	2	US-08-863-028-3	Sequence 3, Appli
11	21	4.2	24	3	US-09-115-061-7	Sequence 7, Appli
12	21	4.2	24	3	US-09-261-079-7	Sequence 7, Appli
13	21	4.2	24	4	US-09-631-349A-1	Sequence 1, Appli
14	21	4.2	28	2	US-08-547-214-55	Sequence 55, Appl
15	21	4.2	28	2	US-08-663-823B-55	Sequence 55, Appl
16	21	4.2	28	3	US-08-942-406-55	Sequence 55, Appl
17	21	4.2	28	3	US-09-322-617-55	Sequence 55, Appl
18	21	4.2	28	4	US-09-751-561-55	Sequence 55, Appl
19	21	4.2	28	4	US-09-724-385-55	Sequence 55, Appl
20	21	4.2	28	4	US-09-757-528-55	Sequence 55, Appl
21	20	4.0	24	1	US-08-547-214-1	Sequence 1, Appli
22	20	4.0	24	2	US-08-547-214-15	Sequence 15, Appl
23	20	4.0	24	2	US-08-663-823B-1	Sequence 1, Appli
24	20	4.0	24	2	US-08-663-823B-15	Sequence 15, Appl
25	20	4.0	24	2	US-08-663-823B-60	Sequence 60, Appl
26	20	4.0	24	2	US-08-663-823B-62	Sequence 62, Appl
27	20	4.0	24	3	US-08-874-825-27	Sequence 27, Appl
28	20	4.0	24	3	US-08-874-825-29	Sequence 29, Appl
29	20	4.0	24	3	US-08-874-825-31	Sequence 31, Appl
30	20	4.0	24	3	US-08-874-825-33	Sequence 33, Appl
31	20	4.0	24	3	US-08-874-825-53	Sequence 53, Appl
32	20	4.0	24	3	US-08-874-825-67	Sequence 67, Appl
33	20	4.0	24	3	US-08-874-825-107	Sequence 107, App
34	20	4.0	24	3	US-08-874-825-109	Sequence 109, App
35	20	4.0	24	3	US-08-663-824-27	Sequence 27, Appl
36	20	4.0	24	3	US-08-663-824-29	Sequence 29, Appl
37	20	4.0	24	3	US-08-663-824-31	Sequence 31, Appl
38	20	4.0	24	3	US-08-663-824-33	Sequence 33, Appl
39	20	4.0	24	3	US-08-663-824-53	Sequence 53, Appl
40	20	4.0	24	3	US-08-663-824-67	Sequence 67, Appl
41	20	4.0	24	3	US-08-663-824-107	Sequence 107, App
42	20	4.0	24	3	US-08-663-824-109	Sequence 109, App
43	20	4.0	24	3	US-08-942-406-1	Sequence 1, Appli
44	20	4.0	24	3	US-08-942-406-15	Sequence 15, Appl
45	20	4.0	24	3	US-09-381-779-3	Sequence 3, Appli
46	20	4.0	24	3	US-09-322-617-1	Sequence 1, Appli
47	20	4.0	24	3	US-09-203-231B-1	Sequence 1, Appli
48	20	4.0	24	3	US-09-203-231B-5	Sequence 5, Appli
49	20	4.0	24	4	US-09-203-231B-19	Sequence 19, Appl
50	20	4.0	24	4	US-09-231-303-27	Sequence 27, Appl
51	20	4.0	24	4	US-09-231-303-29	Sequence 29, Appl
52	20	4.0	24	4	US-09-231-303-31	Sequence 31, Appl
53	20	4.0	24	4	US-09-231-303-33	Sequence 33, Appl
54	20	4.0	24	4	US-09-231-303-53	Sequence 53, Appl
55	20	4.0	24	4	US-09-231-303-67	Sequence 67, Appl
56	20	4.0	24	4	US-09-231-303-107	Sequence 107, App
57	20	4.0	24	4	US-09-231-303-109	Sequence 109, App
58	20	4.0	24	4	US-09-751-561-1	Sequence 1, Appli
59	20	4.0	24	4	US-09-751-561-15	Sequence 15, Appl
60	20	4.0	24	4	US-09-724-385-1	Sequence 1, Appli
61	20	4.0	24	4	US-09-724-385-15	Sequence 15, Appl
62	20	4.0	24	4	US-09-757-528-1	Sequence 1, Appli
63	20	4.0	24	4	US-09-757-528-15	Sequence 15, Appl
64	20	4.0	48	2	US-08-547-214-57	Sequence 57, Appl
65	20	4.0	48	2	US-08-663-823B-57	Sequence 57, Appl
66	20	4.0	48	3	US-08-942-406-57	Sequence 57, Appl
67	20	4.0	48	3	US-09-322-617-57	Sequence 57, Appl
68	20	4.0	48	4	US-09-751-561-57	Sequence 57, Appl
69	20	4.0	48	4	US-09-724-385-57	Sequence 57, Appl
70	20	4.0	48	4	US-09-757-528-57	Sequence 57, Appl
71	20	4.0	53	2	US-08-547-214-59	Sequence 59, Appl
72	20	4.0	53	2	US-08-663-823B-59	Sequence 59, Appl
73	20	4.0	53	3	US-08-942-406-59	Sequence 59, Appl
74	20	4.0	53	3	US-09-322-617-59	Sequence 59, Appl
75	20	4.0	53	4	US-09-751-561-59	Sequence 59, Appl
76	20	4.0	53	4	US-09-724-385-59	Sequence 59, Appl
77	20	4.0	53	4	US-09-757-528-59	Sequence 59, Appl
78	19	3.8	19	1	US-07-974-447-1	Sequence 1, Appli
79	19	3.8	19	1	US-08-149-199-1	Sequence 1, Appli
80	19	3.8	19	2	US-08-381-572-13	Sequence 13, Appl
81	19	3.8	19	2	US-08-592-820-13	Sequence 13, Appl
82	19	3.8	19	2	US-08-993-118-11	Sequence 11, Appl
83	19	3.8	19	3	US-08-748-130-23	Sequence 23, Appl
84	19	3.8	19	3	US-08-845-528C-11	Sequence 11, Appl
85	19	3.8	19	3	US-09-115-061-1	Sequence 1, Appli
86	19	3.8	19	3	US-08-969-815-6	Sequence 6, Appli
87	19	3.8	19	3	US-09-120-025-6	Sequence 6, Appli
88	19	3.8	19	3	US-09-178-973B-1	Sequence 1, Appli
89	19	3.8	19	3	US-09-261-079-1	Sequence 1, Appli
90	19	3.8	19	4	US-09-419-568F-1	Sequence 1, Appli
91	19	3.8	19	4	US-09-354-243B-1	Sequence 1, Appli
92	19	3.8	19	4	US-09-710-481-6	Sequence 6, Appli
93	19	3.8	19	4	US-09-656-952-7	Sequence 7, Appli
94	19	3.8	19	4	US-08-469-260A-1	Sequence 1, Appli
95	19	3.8	19	4	US-09-066-281B-11	Sequence 11, Appl
96	19	3.8	19	4	US-09-553-875-6	Sequence 6, Appli
97	19	3.8	19	4	US-08-488-446-1	Sequence 1, Appli
98	19	3.8	19	4	US-09-768-670-6	Sequence 6, Appli
99	19	3.8	19	4	US-08-467-344A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-631-349A-19
; Sequence 19, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
;   APPLICANT: Abbott Laboratories
;   APPLICANT: Birkenmeyer, Larry G.
;   APPLICANT: Leary, Thomas P.
;   APPLICANT: Muerhoff, A. Scott
;   APPLICANT: Desai, Suresh M.
;   APPLICANT: Mushahwar, Isa K.
;   TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
;   TITLE OF INVENTION: HYBRIDIZATION
;   FILE REFERENCE: 6714.US.O1
;   CURRENT APPLICATION NUMBER: US/09/631,349A
;   CURRENT FILING DATE: 2000-08-02
;   NUMBER OF SEQ ID NOS: 21
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer R-Bam 24
;   NAME/KEY: misc_feature
;   LOCATION: (29)...(32)
;   OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
;   OTHER INFORMATION: positions 29-32
US-09-631-349A-19

Query Match   4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTCTCCAGCCTCTCACCGAGGAT 24
              ||||||||||||||||||||||||
Db   4 ACTCTCCAGCCTCTCACCGAGGAT 27

RESULT 2
US-09-631-349A-19/c
; Sequence 19, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
;   APPLICANT: Abbott Laboratories
;   APPLICANT: Birkenmeyer, Larry G.
;   APPLICANT: Leary, Thomas P.
;   APPLICANT: Muerhoff, A. Scott
;   APPLICANT: Desai, Suresh M.
;   APPLICANT: Mushahwar, Isa K.
;   TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
;   TITLE OF INVENTION: HYBRIDIZATION
;   FILE REFERENCE: 6714.US.O1
;   CURRENT APPLICATION NUMBER: US/09/631,349A
;   CURRENT FILING DATE: 2000-08-02
;   NUMBER OF SEQ ID NOS: 21
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer R-Bam 24
;   NAME/KEY: misc_feature
;   LOCATION: (29)...(32)
;   OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
;   OTHER INFORMATION: positions 29-32
US-09-631-349A-19

Query Match   4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006; Mismatches   24; Conservative   0;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTCTCCAGCCTCTCACCGAGGAT 24
              ||||||||||||||||||||||||
Db   57 ACTCTCCAGCCTCTCACCGAGGAT 34

RESULT 3
US-09-631-349A-21
; Sequence 21, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
;   APPLICANT: Abbott Laboratories
;   APPLICANT: Birkenmeyer, Larry G.
;   APPLICANT: Leary, Thomas P.
;   APPLICANT: Muerhoff, A. Scott
;   APPLICANT: Desai, Suresh M.
;   APPLICANT: Mushahwar, Isa K.
;   TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
;   TITLE OF INVENTION: HYBRIDIZATION
;   FILE REFERENCE: 6714.US.O1
;   CURRENT APPLICATION NUMBER: US/09/631,349A
;   CURRENT FILING DATE: 2000-08-02
;   NUMBER OF SEQ ID NOS: 21
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer R-Bam 19N
;   NAME/KEY: misc_feature
;   LOCATION: (29)...(32)
;   OTHER INFORMATION: n = a or g or c or t/u, unknown or other at
;   OTHER INFORMATION: positions 29-32
US-09-631-349A-21

Query Match   4.8%; Score 24; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 ACTCTCCAGCCTCTCACCGAGGAT 24
              ||||||||||||||||||||||||
Db   4 ACTCTCCAGCCTCTCACCGAGGAT 27

RESULT 4
US-09-631-349A-21/c
; Sequence 21, Application US/09631349A
; Patent No. 6455255
; GENERAL INFORMATION:
;   APPLICANT: Abbott Laboratories
;   APPLICANT: Birkenmeyer, Larry G.
;   APPLICANT: Leary, Thomas P.
;   APPLICANT: Muerhoff, A. Scott
;   APPLICANT: Desai, Suresh M.
;   APPLICANT: Mushahwar, Isa K.
;   TITLE OF INVENTION: METHOD OF PERFORMING SUBTRACTIVE
;   TITLE OF INVENTION: HYBRIDIZATION
;   FILE REFERENCE: 6714.US.O1
;   CURRENT APPLICATION NUMBER: US/09/631,349A
;   CURRENT FILING DATE: 2000-08-02
;   NUMBER OF SEQ ID NOS: 21
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer R-Bam 19N

OM protein - nucleic search, using frame_plus_p2n model

Copyright (c) 1993 - 2003 Compugen Ltd.
GenCore version 5.1.6

Run on: October 23, 2003, 12:36:31 ; Search time 2864 Seconds
(without alignments)
1771.227 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 124
Sequence: 1 LSSLSPRMKSAREVVAVGGK........KAGEILIELDPFAGGVDVAT 124

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2012964

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09484577/runat_23102003_111504_18499/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.lrg -MINMATCH=0.1 -LCOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=250
-USER=US09484577 @CGN_1_1_3508 @runat_23102003_111504_18499 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	8	6.5	231	8	AF358762	AF358762 Oryza sat
c	2	7	5.6	33	6	AR035446	AR035446 Sequence
c	3	7	5.6	42	6	AR091564	AR091564 Sequence
c	4	7	5.6	42	6	AR221814	AR221814 Sequence
c	5	7	5.6	42	6	BD134792	BD134792 Method an
c	6	7	5.6	43	6	AR091560	AR091560 Sequence
c	7	7	5.6	43	6	AR221810	AR221810 Sequence
c	8	7	5.6	43	6	BD134788	BD134788 Method an
	9	7	5.6	45	6	AR035459	AR035459 Sequence
	10	7	5.6	49	6	AR091563	AR091563 Sequence
c	11	7	5.6	49	6	AR091569	AR091569 Sequence
c	12	7	5.6	49	6	AR221813	AR221813 Sequence
c	13	7	5.6	49	6	AR221819	AR221819 Sequence
c	14	7	5.6	49	6	BD134791	BD134791 Method an
c	15	7	5.6	49	6	BD134797	BD134797 Method an
	16	7	5.6	50	6	AR091559	AR091559 Sequence
	17	7	5.6	50	6	AR221809	AR221809 Sequence
	18	7	5.6	50	6	BD134787	BD134787 Method an
	19	7	5.6	56	6	AR091568	AR091568 Sequence
	20	7	5.6	56	6	AR221818	AR221818 Sequence
	21	7	5.6	56	6	BD134796	BD134796 Method an
	22	7	5.6	59	6	AR035460	AR035460 Sequence
	23	7	5.6	59	6	AR035461	AR035461 Sequence
	24	7	5.6	60	6	AR231929	AR231929 Sequence
c	25	7	5.6	60	6	AR231929	AR231929 Sequence
	26	7	5.6	60	6	AR231931	AR231931 Sequence
c	27	7	5.6	60	6	AR231931	AR231931 Sequence
c	28	7	5.6	60	6	AX464727	AX464727 Sequence 27
c	29	7	5.6	60	6	AX464727	AX464727 Sequence 27
	30	7	5.6	60	6	AX464729	AX464729 Sequence
c	31	7	5.6	60	6	AX464729	AX464729 Sequence
	32	7	5.6	70	6	AR035462	AR035462 Sequence
	33	7	5.6	70	6	AR035465	AR035465 Sequence
	34	7	5.6	72	6	AR055681	AR055681 Sequence
	35	7	5.6	76	6	AX240936	AX240936 Sequence
	36	7	5.6	78	6	AR140750	AR140750 Sequence
	37	7	5.6	78	6	I93429	I93429 Sequence 27
	38	7	5.6	78	6	I95056	I95056 Sequence 27
c	39	7	5.6	80	14	HIVLTRCAT	M69075 Human immun
c	40	7	5.6	86	6	AR054880	AR054880 Sequence
c	41	7	5.6	86	6	AR066145	AR066145 Sequence
	42	7	5.6	116	6	AR140738	AR140738 Sequence
	43	7	5.6	116	6	AR140746	AR140746 Sequence
	44	7	5.6	116	6	AR140747	AR140747 Sequence
	45	7	5.6	116	6	I93417	I93417 Sequence 15
	46	7	5.6	116	6	I93425	I93425 Sequence 23
	47	7	5.6	116	6	I93426	I93426 Sequence 24
	48	7	5.6	116	6	I95044	I95044 Sequence 15
	49	7	5.6	116	6	I95052	I95052 Sequence 23
	50	7	5.6	116	6	I95053	I95053 Sequence 24

51	7	125	6	AR140737	5.6	AR140737 Sequence
52	7	125	6	I93416	5.6	I93416 Sequence 14
53	7	125	6	I95043	5.6	I95043 Sequence 14
54	7	128	14	HIVIU15071	5.6	U15071 Human immun
55 c	7	129	6	AX526985	5.6	AX526985 Sequence
56 c	7	141	14	HIVIU15068	5.6	U15068 Human immun
57	7	159	6	BD028280	5.6	BD028280 Sequence
58 c	7	168	12	S66479	5.6	S66479 orn=ornati
59	7	182	8	AF525491	5.6	AFS25491 Talaromyc
60 c	7	186	14	AF272007	5.6	AF272007 HIV-1 iso
61 c	7	187	8	AY200846	5.6	AY200846 Arabidops
62	7	189	6	AR029373	5.6	AR029373 Sequence
63	7	197	6	AAU33639	5.6	U33639 Apteronotus
64 c	7	197	5	ALU33649	5.6	U33649 Apteronotus
65	7	197	14	AF272009	5.6	AF272009 HIV-1 iso
66	7	203	14	REHTLV3B	5.6	X05821 Human T-lym
67 c	7	210	6	AX523415	5.6	AX523415 Sequence
68	7	218	6	A64094	5.6	A64094 Sequence 20
69	7	218	8	HVU234711	5.6	AJ234711 Hordeum v
70 c	7	220	6	BD169510	5.6	BD169510 T cell li
71	7	220	14	HIVNE020	5.6	M58192 Human immun
72	7	220	14	HIVNE039	5.6	M58211 Human immun
73 c	7	224	10	MUSFHCA	5.6	M73678 Mus musculu
74	7	224	14	AF000533	5.6	AF000533 HIV-1 str
75	7	224	14	AF000538	5.6	AF000538 HIV-1 str
76	7	224	14	AF000543	5.6	AF000543 HIV-1 str
77 c	7	231	8	AY065587	5.6	AY065587 Zea mays
78 c	7	231	8	AY065589	5.6	AY065589 Zea mays
79	7	240	6	A03792	5.6	A03792 Nucleotide
80	7	240	6	A04709	5.6	A04709 HIV-1 probe
81	7	240	6	A14070	5.6	A14070 HIV-1 RNA.
82 c	7	240	8	PPO19468	5.6	Y19468 Phaseolus p
83	7	240	14	AY221712	5.6	AY221712 HIV-1 iso
84	7	242	14	AF264864	5.6	AF264864 Nothomyrm
85	7	242	6	AR001163	5.6	AR001163 Sequence
86	7	242	6	AR003041	5.6	AR003041 Sequence
87	7	242	6	AR033015	5.6	AR033015 Sequence
88	7	242	6	AR209679	5.6	AR209679 Sequence
89	7	242	6	I29755	5.6	I29755 Sequence 62
90	7	242	6	I76885	5.6	I76885 Sequence 27
91	7	242	6	I87837	5.6	I87837 Sequence 27
92	7	242	6	I91429	5.6	I91429 Sequence 62
93	7	243	14	HIVNE012	5.6	M58184 Human immun
94	7	244	14	ABC12645	5.6	AB012645 Homo sapi
95	7	244	9	AF000535	5.6	AF000535 HIV-1 str
96	7	245	6	BD081963	5.6	BD081963 Method. 8
97	7	249	6	AX023665	5.6	AX023665 Sequence
98 c	7	249	6	AX655508	5.6	AX655508 Sequence
99 c	7	249	14	HCVNS52	5.6	Z36526 Hepatitis C
100	6	20	6	AR299900	4.8	AR299900 Sequence

ALIGNMENTS

RESULT 1
AF358762/c

```
LOCUS       AF358762           231 bp    mRNA    linear   PLN 24-JAN-2002
DEFINITION  Oryza sativa clone Osh67 unknown protein mRNA, partial cds.
ACCESSION   AF358762
VERSION     AF358762.1  GI:18308221
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 231)
  AUTHORS   Lee,R.H., Wang,C.H., Huang,L.T. and Chen,S.C.
  TITLE     Leaf senescence in rice plants: cloning and characterization of
            senescence up-regulated genes
  JOURNAL   J. Exp. Bot. 52 (358), 1117-1121 (2001)
   MEDLINE  21336245
   PUBMED   11443928
REFERENCE   2  (bases 1 to 231)
  AUTHORS   Lee,R.H. and Chen,S.C.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2001) Institute of Botany, Academia Sinica,
            Taipei 11529, Taiwan
FEATURES             Location/Qualifiers
     source          1..231
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /cultivar="Tainong 57"
                     /db_xref="taxon:4530"
                     /clone="Osh67"
                     /tissue_type="dark-induced senescent leaf"
                     /dev_stage="two-week old seedling"
     CDS             <1..>231
                     /codon_start=3
                     /product="unknown protein"
                     /protein_id="AAL67875.1"
                     /db_xref="GI:18308222"
                     /translation="TSPQFPMLQESVLQSFKLPRAVLCIGGVVKIEPLGRVQKQEMDD
                     LYTICISHVQIVGIPLPRELGVDPYKNGVVLK"
BASE COUNT       62 a     48 c     56 g     65 t
ORIGIN
```

Alignment Scores:
Pred. No.:   86.6   Length: 231
Score: 8.00   Matches: 8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 6.45%   Indels: 0
DB: 8   Gaps: 0

US-09-484-577A-4 (1-124) x AF358762 (1-231)

```
Qy  29 ProAlaAlaLeuGluIleValGlu 36
       ||||||||||||||||||||||||
Db  30 CCTGCAGCATTGGAAATTGTCGAG 7
```

RESULT 2
AR035446

```
LOCUS       AR035446            33 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 18 from patent US 5871902.
ACCESSION   AR035446
VERSION     AR035446.1  GI:5952114
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Weininger,S. and Weininger,A.M.
  TITLE     Sequence-specific detection of nucleic acid hybrids using a
            DNA-binding molecule or assembly capable of discriminating perfect
            hybrids from non-perfect hybrids
            Patent: US 5871902-A 18 16-FEB-1999;
FEATURES             Location/Qualifiers
     source          1..33
                     /organism="unknown"
BASE COUNT        2 a      9 c     17 g      5 t
ORIGIN
```

Alignment Scores:
Pred. No.:   141   Length: 33
Score: 7.00   Matches: 7
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.65%   Indels: 0
DB: 6   Gaps: 0

US-09-484-577A-4 (1-124) x AR035446 (1-33)

```
Qy  58 ValAlaTrpAlaGlyLeuGly 64
       |||||||||||||||||||||
Db   2 GTGGCCTGGCGCGGACTGGGG 22
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - nucleic search, using frame_plus_p2n model

Run on: October 23, 2003, 13:28:22 ; Search time 1724 Seconds
(without alignments)
1748.118 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 124
Sequence: 1 LSSLSPRMKSAREVVAVGGK.........KAGEILIELDPFAGGVDVAT 124

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4672106

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09484577/runat_23102003_111504_18512/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=250
-USER=US09484577 @CGN_1_1_2810 @runat_23102003_111504_18512 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_eston:*
18: em_gss_hum:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Query Length	DB	ID	Description
c	1	9	7.3	180	9	AJ477791	AJ477791 AJ477791
c	2	9	7.3	225	9	AU082059	AU082059 AU082059
c	3	9	7.3	249	9	AI597979	AI597979 ts04h05.x
	4	8	6.5	205	12	BM594596	BM594596 170006688
	5	8	6.5	225	12	AZ578485	AZ578485 23h12 Sho
c	6	8	6.5	244	12	BM876055	BM876055 if87b01.y
	7	7	5.6	29	28	AZ853380	AZ853380 2M0156J03
c	8	7	5.6	49	9	AA113042	AA113042 zn65d12.r
	9	7	5.6	64	9	AA894355	AA894355 oj16g12.s
c	10	7	5.6	72	10	BF633486	BF633486 NF055B12D
	11	7	5.6	76	13	BQ094017	BQ094017 fz54e04.y
	12	7	5.6	78	29	AL938849	AL938849 Arabidops
	13	7	5.6	83	14	BI082225	BI082225 60287403
	14	7	5.6	89	14	R65877	R65877 yi23d03.r1
	15	7	5.6	95	13	BU652348	BU652348 1112099D0
	16	7	5.6	102	14	BG953925	BG953925 RC4-CT065
c	17	7	5.6	104	14	CA453170	CA453170 EST01 Hum
c	18	7	5.6	105	9	AA839373	AA839373 vw48g05.r
	19	7	5.6	106	9	AL503786	AL503786 AL503786
c	20	7	5.6	107	12	BI052864	BI052864 RC0-GN027
	21	7	5.6	110	9	AJ503237	AJ503237 AJ503237
c	22	7	5.6	112	14	BG658097	BG658097 TgESTzya4
c	23	7	5.6	112	14	CA044299	CA044299 ssalplnb5
	24	7	5.6	112	14	CB504908	CB504908 ssalplnb5
c	25	7	5.6	113	9	AW739688	AW739688 BRI00201
c	26	7	5.6	113	9	AW739787	AW739787 BRI00300
	27	7	5.6	123	12	BM884009	BM884009 rc18c12.y
	28	7	5.6	124	14	BF414974	BF414974 UI-R-BJ2-
	29	7	5.6	124	14	R63432	R63432 yi08e05.r1
c	30	7	5.6	124	28	BH620760	BH620760 1007098C0
c	31	7	5.6	125	28	AZ920101	AZ920101 1006018B0
	32	7	5.6	125	28	BH849615	BH849615 SALK_0699
	33	7	5.6	127	12	BI396580	BI396580 ro54h01.y
c	34	7	5.6	128	12	AZ038470	AZ038470 RPCI-23-3
c	35	7	5.6	130	12	BM738329	BM738329 K-EST0003
c	36	7	5.6	130	13	BY288555	BY288555 BY288555
	37	7	5.6	130	29	CC335874	CC335874 OGPBB21TH
c	38	7	5.6	131	9	AF091519	AF091519 AF091519
	39	7	5.6	131	14	CB503345	CB503345 ssalmges0
	40	7	5.6	131	12	BH628286	BH628286 1007077EC
	41	7	5.6	132	12	BM040114	BM040114 TgESTzya5
	42	7	5.6	132	13	BY144970	BY144970 BY144970
	43	7	5.6	132	13	BH124409	BH124409 RPCI-24-3
	44	7	5.6	133	14	CA778843	CA778843 MPL384_11
	45	7	5.6	135	13	BY129163	BY129163 BY129163
c	46	7	5.6	135	28	BZ784047	BZ784047 PUFJN36TB
	47	7	5.6	136	28	BH667508	BH667508 BOMBD71TF
	48	7	5.6	138	9	AF091520	AF091520 AF091520
c	49	7	5.6	139	12	BE500523	BE500523 WHE0987-0
	50	7	5.6	139	12	BI033900	BI033900 QV2-NN200
	51	7	5.6	140	9	AV131901	AV131901 AV131901
	52	7	5.6	142	9	AI735737	AI735737 at2g08.x
c	53	7	5.6	142	12	BI033901	BI033901 QV2-NN200
	54	7	5.6	143	10	BB571857	BB571857 BB571857
	55	7	5.6	143	10	BB700540	BB700540 QV3-NN010
c	56	7	5.6	143	12	BI679559	BI679559 949003E09
	57	7	5.6	145	13	BY136025	BY136025 BY136025
	58	7	5.6	145	10	AA013190	AA013190 ze28e04.s
c	59	7	5.6	145	10	BB589065	BB589065 BB589065
c	60	7	5.6	145	28	AZ581119	AZ581119 1M0369K23
c	61	7	5.6	148	13	BU208741	BU208741 603951712
c	62	7	5.6	148	28	AZ500752	AZ500752 1M0339J18

```
63    7   149  13  BQ835854         BQ835854 rf58g02.y
64    7   150   9  AW407072         AW407072 UI-HF-BL0
c  65    7   155  14  CD017371         CD017371 NXCI_110
c  66    7   156  10  BB581435         BB581435 BB581435_
67    7   157   9  AU255443         AU255443 AU255443
c  68    7   157   9  AV017222         AV017222 AV017222
69    7   160  13  BQ565096         BQ565096 gi30e12.y
c  70    7   160  29  BZ768706         BZ768706 SALK_1406
71    7   163  13  BY369814         BY369814 BY369814
c  72    7   164  10  BE700464         BE700464 QV3-NN010
73    7   165  10  HSM095483        Bx501766 Homo sapi
c  74    7   165  10  BB546025         BB546025 BB546025
75    7   165  13  BY278133         BY278133 BY278133
c  76    7   165  28  AQ314942         AQ314942 RPCII1-95
77    7   166   9  AA673042         AA673042 vn7igl1.r
78    7   167  10  BF813858         BF813858 PM4-C1011
79    7   168   9  AL043724         AL043724 DKFZp434L
80    7   168  12  BM127195         BM127195 ie95908.x
81    7   168  14  CD151514         CD151514 SSH64 Tam
82    7   170  10  BF915692         BF915692 MR3-JT012
83    7   170  14  T07162           T07162 EST05051 Fe
84    7   170  28  AQ184073         AQ184073 HS_3216_A
c  85    7   171   9  AA876993         AA876993 ny49f07.s
86    7   171   9  AW114718         AW114718 rs53g11.y
87    7   171  29  BZ737358         BZ737358 OGECX74TM
88    7   173  10  BF841452         BF841452 RC2-HT107
89    7   174   9  AA876752         AA876752 rz48a01.s
90    7   174   9  AA886051         AA886051 ny44c05.s
91    7   174  14  CA998983         CA998983 S34S_H03
c  92    7   174  28  B78965           B78965 CIT-HSP-493
c  93    7   174  28  B88275           B88275 CIT-HSP-493
94    7   175   9  AA876995         AA876995 ny49c09.s
95    7   175   9  AA586327         AA586327 rd18b2.s
96    7   176  28  AZ463283         AZ463283 1MO272EC1
97    7   178  10  BE045502         BE045502 hk77d03.x
c  98    7   179  12  BI023108         BI023108 MR3-YT032
c  99    7   180   9  AJ463996         AJ463996 AJ463996
c 100    7   181   9  AA658468         AA658468 ru21h2.s
```

ALIGNMENTS

```
RESULT 1
AJ477791/c
LOCUS       AJ477791         180 bp    mRNA      linear   EST 24-MAY-2002
DEFINITION  AJ477791 S00010 Hordeum vulgare cDNA clone S00C100000SGC8F1, mRNA
            sequence.
ACCESSION   AJ477791
VERSION     AJ477791.  GI:2119746
KEYWORDS    EST.
SOURCE      Hordeum vulgare
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 180)
  AUTHORS   Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
  TITLE     Barley EST's
  JOURNAL   Unpublished
  COMMENT   Contact: Schulman AH
            Institute of Biotechnology
            University of Helsinki
            P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
            Finland.
FEATURES             Location/Qualifiers
     source          1..180
                     /organism="Hordeum vulgare"
                     /mol_type="mRNA"
                     /db_xref="taxon:4513"
                     /clone="S0001000005G08F1"
                     /tissue_type="Developing seed"
                     /dev_stage="0.,3-,6-,9-days after pollination"
                     /clone_lib="S00010"
```

BASE COUNT   37 a   54 c   57 g   32 t
ORIGIN

Alignment Scores:
Pred. No.:   230   Length: 180
Score: 9.00   Matches: 9
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 7.26%   Indels: 0
DB: 9   Gaps: 0

US-09-484-577A-4 (1-124) x AJ477791 (1-180)

Qy   22 ArgAspGluLeuAlaPheLeuProAla 30
           |||||||||||||||||||||||||||
Db   46 CGCGACGAACTTGCCTTCCCTCCCTGCT 20

```
RESULT 2
AU082059/c
LOCUS       AU082059         225 bp    mRNA      linear   EST 02-APR-2002
DEFINITION  AU082059 Rice panicle at flowering stage Oryza sativa (japonica
            cultivar-group) cDNA clone E0309, mRNA sequence.
ACCESSION   AU082059
VERSION     AU082059.1  GI:6536630
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 225)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from panicle at flowering stage
  JOURNAL   Unpublished
  COMMENT   Contact: Takuji Sasaki
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'
FEATURES             Location/Qualifiers
     source          1..225
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="E0309"
                     /dev_stage="flowering stage"
                     /clone_lib="Rice panicle at flowering stage"
                     /note="Organ: panicle; Rice cDNA from panicle at flowering
                     stage"
```

BASE COUNT   28 a   89 c   67 g   41 t
ORIGIN

Alignment Scores:
Pred. No.:   281   Length: 225
Score: 9.00   Matches: 9
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 7.26%   Indels: 0
DB: 9   Gaps: 0

US-09-484-577A-4 (1-124) x AU082059 (1-225)

Qy   85 ValGlnProLeuGluValGlyValVal 93
           |||||||||||||||||||||||||||
Db   71 GTGCAGCCACTCGAAGTAGGGGTCGTC 45

```
RESULT 3
AI597979/c
LOCUS       AI597979         249 bp    mRNA      linear   EST 14-MAY-1999
```

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n_model

Run on: October 23, 2003, 12:35:16 ; Search time 235 Seconds
                                           (without alignments)
                                           1424.384 Million cell updates/sec

Title: US-09-484-577A-4
Perfect score: 124
Sequence: 1 LSSLSPRMKSAREVVAVGGK.......KAGEILIELDPFAGGVDVAT 124

Scoring table: OLIGO
               Xgapop 60.0 , Xgapext 60.0
               Ygapop 60.0 , Ygapext 60.0
               Fgapop 6.0 , Fgapext 7.0
               Delop   6.0 , Delext   7.0

Searched: 252756 seqs, 134971017 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3106038

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
Q=/cgn2_1/USPTO_spool/US09484577//runat_23102003_l11503_18487/app_query.fasta_1.263
DB=N_Geneseq_19Jun03 -QFMT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
LIST=100 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=75 -MODE=LOCAL
OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=250
-USER=US09484577 @CGN_1_1_0 @runat_23102003_l11503_18487 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   N_Geneseq_19Jun03:*
       1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
       2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA:981.DAT:*
       3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA:982.DAT:*
       4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA:983.DAT:*
       5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA:984.DAT:*
       6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
       7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA:986.DAT:*
       8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
       9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
      10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
      11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
      12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
      13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
      14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
      15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
      16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
      17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
      18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
      19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
      20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
      21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
      22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
      23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
      24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
      25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		7	5.6	33	17	AAT30598	Target binding reg
2		7	5.6	36	14	AAQ36247	HIV-1 LTR region c
3		7	5.6	38	13	AAQ33271	HIV-1 LTR - corres
4		7	5.6	43	12	AAQ12253	HIV-1 LTR mutation
5	c	7	5.6	43	22	AAC86352	EcoRI-AlwNI linker
6		7	5.6	45	17	AAT30611	Target binding reg
7		7	5.6	47	24	AAD29073	bxII gene 5' end c
8		7	5.6	48	12	AAQ12252	HIV-1 LTR mutation
9		7	5.6	48	12	AAQ14775	Rennin gene probe.
10	c	7	5.6	49	22	AAC86358	Linker used to mak
11		7	5.6	50	22	AAC86351	EcoRI-AlwNI linker
12		7	5.6	56	22	AAC86357	Linker used to mak
13		7	5.6	59	17	AAT30612	Target binding reg
14		7	5.6	59	17	AAT30613	Target binding reg
15		7	5.6	60	21	AAA15068	Oligonucleotide us
16		7	5.6	60	21	AAD00174	Oligo BET-039 to g
17		7	5.6	60	24	AAD32135	Sau3AI DNA fragmen
18	c	7	5.6	60	24	AAD32137	Sau3AI DNA fragmen
19		7	5.6	60	24	AAD32137	Sau3AI DNA fragmen
20	c	7	5.6	60	24	AAD32137	Sau3AI DNA fragmen
21	c	7	5.6	70	17	AAT30614	Target binding reg
22		7	5.6	70	17	AAT30617	Probe nucleic acid
23		7	5.6	72	12	AAQ12254	HIV-1 Spi-bind
24		7	5.6	72	20	AAT06609	HIV wild type pILI
25		7	5.6	72	22	AAV34381	Wild type pILIC LT
26		7	5.6	76	22	AAS23493	C. albicans essent
27		7	5.6	78	18	AAT65169	Transforming growt
28	c	7	5.6	86	18	AAT78708	Class 2 SELEX gene
29		7	5.6	104	22	ABA50342	Human breast cell
30		7	5.6	104	22	ABA68293	Human foetal liver
31		7	5.6	104	22	ABA35295	Probe #13761 for g
32		7	5.6	104	22	AAK16665	Human brain expres
33		7	5.6	104	22	AAI23184	Probe #13117 for g
34		7	5.6	104	22	AAI48507	Probe #17193 used
35		7	5.6	104	22	AAI08837	Probe #8828 used t
36		7	5.6	104	23	ABS42038	Human liver single
37		7	5.6	104	24	ABS16483	Human genome-deriv
38		7	5.6	116	18	AAT65157	Transforming growt
39		7	5.6	116	18	AAT65165	Transforming growt
40		7	5.6	116	18	AAT65166	Transforming growt
41		7	5.6	125	18	AAT65156	Transforming growt
42		7	5.6	129	22	AAS30475	DNA encoding novel
43		7	5.6	129	22	AAS30476	DNA encoding novel
44		7	5.6	129	22	AAL06255	Human reproductive
45		7	5.6	129	22	AAL06256	Human reproductive
46		7	5.6	129	22	AAK90424	Human digestive sy
47		7	5.6	129	22	AAK90425	Human digestive sy
48		7	5.6	129	22	AAI27999	Probe #17932 for g
49	c	7	5.6	129	20	ABN97468	Human NEDD-1 exon
50		7	5.6	153	20	ABV81498	Nucleic acid compr
51	c	7	5.6	159	21	AAC04535	Human secreted pro
52		7	5.6	167	24	ABL72791	Corn tassel-derive
53		7	5.6	168	12	AAQ13326	OrnatinE gene. Pl
54	c	7	5.6	169	21	AAF14750	Aspergillus oryzae
55	c	7	5.6	174	24	ABA92779	HIV-1 promoter seq
56	c	7	5.6	178	22	ABA07490	Human ovarian and
57	c	7	5.6	178	22	AAL00412	Human reproductive
58	c	7	5.6	189	17	AAT07026	Immunogen DNA from
59	c	7	5.6	189	18	AAT67569	Coding sequence fo
60	c	7	5.6	190	25	ABX24072	Human GDP-mannose
61	c	7	5.6	200	19	AAV07529	Human GDP-mannose
62	c	7	5.6	210	24	ABQ75264	HIV-1 LTR partial
63	c	7	5.6	210	25	ABX27484	Human lung specifi
64	c	7	5.6	216	12	AAQ13832	Human GDP-mannose
65	c	7	5.6	218	18	AAT91145	Strawberry fruit r
66	c	7	5.6	220	24	ABL60603	HIV long terminal

c	67	7	5.6	228	25	ABX26565	Human GDP-mannose
	68	7	5.6	235	24	ABN22177	Human ORFX polynuc
	69	7	5.6	235	25	ABX49799	Bovine EST associa
	70	7	5.6	240	21	AAZ89627	HIV-1 3'-end DNA f
	71	7	5.6	242	15	AAQ72878	HIV-1 proviral pro
	72	7	5.6	242	18	AAT49561	HIV LTR sequence.
	73	7	5.6	242	20	AAX17627	HIV proviral promo
	74	7	5.6	242	24	ABK83118	DNA binding molecu
	75	7	5.6	245	20	AAZ11438	HIV derived synthe
c	76	7	5.6	246	16	AAT19601	Human gene signatu
	77	7	5.6	249	21	AAA11999	Murine PGK HRE der
c	78	6	4.8	19	21	AAD40414	Bovine DGAT1 gene
	79	6	4.8	20	21	AAZ77279	Human biallelic ma
c	80	6	4.8	20	21	AAA66536	Dog genomic marker
c	81	6	4.8	20	21	AAA66598	Dog genomic marker
c	82	6	4.8	21	18	AAT51136	Chromosomal locati
	83	6	4.8	21	24	ABS97556	Human epoxide hydr
	84	6	4.8	21	24	ABS98398	Human multidrug re
c	85	6	4.8	21	24	ABL42008	Nucleotide sequenc
	86	6	4.8	22	24	ABK92070	Novel secreted pro
	87	6	4.8	22	24	ABZ58846	A'B zinc finger pr
	88	6	4.8	23	17	AAT10334	Anti-P.aeruginosa
	89	6	4.8	23	22	AAH23398	Nucleotide sequenc
	90	6	4.8	23	25	ABA00925	Non-biotinylated p
	91	6	4.8	24	15	AAQ65638	Primer/adaptor use
	92	6	4.8	24	15	AAQ65644	Primer/adaptor use
	93	6	4.8	24	15	AAQ78833	HCMV Ab lambda lig
	94	6	4.8	24	16	AAT00032	HGBV DNA PCR prime
	95	6	4.8	24	17	AAT39259	Primer, Lu3, used
	96	6	4.8	24	17	AAT13195	Primer R Bgl 24 us
	97	6	4.8	24	17	AAT13201	Primer R Bam 24 us
	98	6	4.8	24	18	AAV06189	Primer used when o
c	99	6	4.8	24	18	AAT92397	Adapter primer RA2
	100	6	4.8	24	18	AAT92398	Adapter primer RC2

ALIGNMENTS

RESULT 1
AAT30598
ID   AAT30598 standard; DNA; 33 BP.
XX
AC   AAT30598;
XX
DT   19-FEB-1997 (first entry)
XX
DE   Target binding region #18.
XX
KW   Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SP1;
KW   TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus;
KW   HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence;
KW   nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria;
KW   virus; ss.
XX
OS   Synthetic.
XX
PN   WO9617956-A2.
XX
PD   13-JUN-1996.
XX
PF   07-DEC-1995; 95WO-US15944.
XX
PR   09-DEC-1994; 94US-0353476.
XX
PA   (GENE-) GENE POOL INC.
XX
PI   Weininger AM, Weininger S;
XX
DR   WPI; 1996-287199/29.
XX
PT   Probe nucleic acids, target binding assemblies, etc - for detection
PT   a localisation of specific nucleic acid sequences, esp. HIV and
PT   HPv

XX
PS   Disclosure; Page 69; 172pp; English.
XX
CC   AAT30581-T30614 represent target binding regions (TBR) of a probe of the
CC   invention. The probe of the invention contains a TBR, a booster binding
CC   region (BBR), and an optional support or attachment (OSA). The target
CC   binding assembly (TBA) recognised by the probe, contains at least one
CC   nucleic acid recognition unit (NAR), and optionally a linker sequence,
CC   an assembly sequence (see AAR95994-R95998), an asymmetry sequence (see
CC   AAR95999-R96006), a nuclear localisation signal sequence (see AAR96007),
CC   and an OSA. The assembly sequence and asymmetry sequences are
CC   responsible for the folding and association of the NARs. The NARs (see
CC   AAR95965-R95993) are selected from NF-kappa-B, SP1, TATA, human
CC   papillomavirus (HPV) E2, HPV LTR, human immunodeficiency virus (HIV) LTR
CC   and Tat binding units. The linker sequence is an oligopeptide, which
CC   does not interfere with NAR function, but provides stability and control
CC   over the spacing of the NAR from the rest of the TBA. The OSA is an
CC   attached support or indicator, or other means of localisation of the
CC   probe. The probe can be used in a method for detecting or localising a
CC   specific target nucleic acid sequence (TNA). The method is highly
CC   sensitive, and has a high degree of specificity. The method can be used
CC   for detecting specific nucleic acid sequences, including those found in
CC   human cells, in HIV, HPV, and other nucleic acid containing systems,
CC   including bacteria and viruses.
XX
SQ   Sequence 33 BP; 2 A; 9 C; 17 G; 5 T; 0 other;

Alignment Scores:
Pred. No.:   102   Length: 33
Score: 7.00   Matches: 7
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.65%   Indels: 0
DB: 17   Gaps: 0

US-09-484-577A-4 (1-124) x AAT30598 (1-33)

Qy   58 ValAlaTrpAlaGlyLeuGly 64
          ||||||||||||||||||||||
Db   2 GTGGCCTGGGCGGGACTGGGG 22

RESULT 2
AAQ36247
ID   AAQ36247 standard; DNA; 36 BP.
XX
AC   AAQ36247;
XX
DT   25-MAR-2003 (updated)
DT   07-JUN-1993 (first entry)
XX
DE   HIV-1 LTR region comprising binding site for Sp1-line transcription
DE   activator.
XX
KW   Human immunodeficiency virus; AIDS; transcription initiation;
KW   long terminal repeat; target; duplex; ss.
XX
OS   Synthetic.
XX
PN   US5176996-A.
XX
PD   05-JAN-1993.
XX
PF   22-DEC-1989; 89US-0453532.
XX
PR   20-DEC-1988; 88US-0287359.
PR   22-DEC-1989; 89US-0453532.
XX
PA   (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI   Hogan ME, Kessler DJ;
XX
DR   WPI; 1993-035718/04.
XX

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:    October 23, 2003, 14:16:12 ; Search time 236 Seconds
                                        (without alignments)
                                1409.082 Million cell updates/sec

Title:          US-09-484-577a-4
Perfect score:  124
Sequence:       1 LSSLSPRMKSAREVVAVGGK.........KAGEILIELDPFAGGVDVAT 124

Scoring table:  OLIGO
                Xgapop 60.0 , Xgapext 60.0
                Ygapop 60.0 , Ygapext 60.0
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1792395 seqs, 1340900451 residues

Word size:      1

Total number of hits satisfying chosen parameters:       1463635

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09484577/runat_23102003_111506_18614/app_query.fasta_1.263
-DB=Published Applications NA -QFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=100 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=250 -USER=US09484577 @CGN_1_1_285 @runat_23102003_111506_18614 -NCPU=6
-ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPCP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :      Published Applications NA:*
         1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
         2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
         3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
         4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
         5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
         6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
         7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
         8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
         9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
        10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
        11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
        12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
        13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
        14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
        15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
        16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
        17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

                               SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

Result No.	Score	% Query Match	Length	DB ID	Description
1	7	5.6	33	8	US-08-860-844-18	Sequence 18, Appl

```
  2  c   5.6  7   33  12  US-10-407-543-18   Sequence 18, Appl
  3  c   5.6  7   42  13  US-10-192-085-17   Sequence 17, Appl
  4      5.6  7   43  13  US-10-192-085-13   Sequence 13, Appl
  5      5.6  7   45   8  US-08-860-844-31   Sequence 31, Appl
  6      5.6  7   45  12  US-10-407-543-31   Sequence 31, Appl
  7  c   5.6  7   49  13  US-10-192-085-16   Sequence 16, Appl
  8      5.6  7   49  13  US-10-192-085-22   Sequence 22, Appl
  9      5.6  7   50  13  US-10-192-085-12   Sequence 12, Appl
 10      5.6  7   56  13  US-10-192-085-21   Sequence 21, Appl
 11      5.6  7   59   8  US-08-860-844-32   Sequence 32, Appl
 12      5.6  7   59   8  US-08-860-844-33   Sequence 33, Appl
 13      5.6  7   59  12  US-10-407-543-32   Sequence 32, Appl
 14      5.6  7   59  12  US-10-407-543-33   Sequence 33, Appl
 15      5.6  7   60  10  US-09-832-659-5    Sequence 5, Appli
 16      5.6  7   60  10  US-09-832-658-8    Sequence 8, Appli
 17      5.6  7   70   8  US-08-860-844-34   Sequence 34, Appl
 18      5.6  7   70   8  US-08-860-844-37   Sequence 37, Appl
 19      5.6  7   70  12  US-10-407-543-34   Sequence 34, Appl
 20      5.6  7   70  12  US-10-407-543-37   Sequence 37, Appl
 21      5.6  7   78  12  US-10-223-666-27   Sequence 27, Appl
 22  c   5.6  7  104   9  US-09-864-761-20615  Sequence 20615, A
 23  c   5.6  7  108  12  US-10-029-386-18332  Sequence 18332, A
 24      5.6  7  116  12  US-10-223-666-15   Sequence 15, Appl
 25      5.6  7  116  12  US-10-223-666-23   Sequence 23, Appl
 26      5.6  7  116  12  US-10-223-666-24   Sequence 24, Appl
 27      5.6  7  125  12  US-10-223-666-14   Sequence 14, Appl
 28  c   5.6  7  129   9  US-09-864-761-33275  Sequence 33275, A
 29  c   5.6  7  129  10  US-09-872-462-15   Sequence 15, Appl
 30      5.6  7  129  11  US-09-764-891-8943   Sequence 8943, Ap
 31      5.6  7  129  11  US-09-764-891-8944   Sequence 8944, Ap
 32      5.6  7  138  12  US-10-029-386-21849  Sequence 21849, A
 33  c   5.6  7  167   9  US-09-294-093B-2165  Sequence 2165, Ap
 34  c   5.6  7  178  11  US-09-764-891-413  Sequence 413, App
 35  c   5.6  7  178  14  US-10-205-428-47   Sequence 47, Appl
 36  c   5.6  7  190  10  US-09-878-574-6131   Sequence 6131, Ap
 37  c   5.6  7  210  10  US-09-878-574-9543   Sequence 9543, Ap
 38  c   5.6  7  210  13  US-10-001-857-3    Sequence 3, Appli
 39  c   5.6  7  228  10  US-09-878-574-8624   Sequence 8624, Ap
 40  c   5.6  7  235  10  US-09-960-352-14964  Sequence 14964, A
 41      4.7  7  242  11  US-09-993-346-627  Sequence 627, Appl
 42      4.8  6   19  12  US-10-251-117-809  Sequence 809, App
 43  c   4.8  6   19  12  US-10-251-117-1116   Sequence 1116, Ap
 44      4.8  6   22   9  US-09-772-105-37   Sequence 37, Appl
 45      4.8  6   23   9  US-09-755-398A-2   Sequence 2, Appli
 46      4.8  6   23  11  US-10-191-438-2    Sequence 2, Appli
 47      4.8  6   23  14  US-10-191-438-4    Sequence 4, Appli
 48      4.8  6   24   8  US-08-424-550B-1   Sequence 1, Appli
 49      4.8  6   24   9  US-09-751-561-1    Sequence 1, Appli
 50      4.8  6   24   9  US-09-751-561-15   Sequence 15, Appl
 51      4.8  6   24   9  US-09-751-797-1    Sequence 1, Appli
 52      4.8  6   24  10  US-09-928-457-55   Sequence 55, Appl
 53      4.8  6   24  10  US-09-928-457-58   Sequence 58, Appl
 54      4.8  6   24  11  US-09-989-364-1    Sequence 1, Appli
 55      4.8  6   24  11  US-09-989-364-5    Sequence 5, Appli
 56      4.8  6   24  11  US-09-989-364-19   Sequence 19, Appl
 57      4.8  6   24  11  US-09-930-334-7    Sequence 7, Appli
 58      4.8  6   24  12  US-09-989-534-69   Sequence 69, Appli
 59      4.8  6   24  12  US-10-134-345-7    Sequence 7, Appli
 60      4.8  6   24  12  US-10-348-190-22   Sequence 22, Appl
 61      4.8  6   24  12  US-10-160-237-11   Sequence 11, Appl
 62      4.8  6   24  13  US-10-032-626-7    Sequence 7, Appli
 63      4.8  6   24  13  US-10-067-813-6    Sequence 6, Appli
 64      4.8  6   24  13  US-10-088-108-11   Sequence 11, Appl
 65      4.8  6   24  14  US-10-225-264-4    Sequence 4, Appli
 66      4.8  6   24  14  US-10-193-451A-7   Sequence 7, Appli
 67      4.8  6   24  14  US-10-193-451A-13  Sequence 13, Appl
 68      4.8  6   24  14  US-10-193-451A-19  Sequence 19, Appl
 69      4.8  6   24  14  US-10-193-451A-25  Sequence 25, Appl
 70      4.8  6   24  14  US-10-001-670-27   Sequence 27, Appl
 71      4.8  6   24  14  US-10-001-670-29   Sequence 29, Appl
 72      4.8  6   24  14  US-10-001-670-31   Sequence 31, Appl
 73      4.8  6   24  14  US-10-001-670-33   Sequence 33, Appl
 74      4.8  6   24  14  US-10-001-670-53   Sequence 53, Appl
```

```
 75   6  24  14  4.8  US-10-001-670-67       Sequence 67, Appl
 76   6  24  14  4.8  US-10-001-670-107      Sequence 107, App
 77   6  24  14  4.8  US-10-001-670-109      Sequence 109, App
 78   6  25  14  4.8  US-10-215-112-7793     Sequence 7793, Ap
 79   6  25  14  4.8  US-10-215-112-7919     Sequence 7919, Ap
c80   6  25  14  4.8  US-10-098-263B-31415   Sequence 31415, A
 81   6  25  14  4.8  US-10-098-263B-35846   Sequence 35846, A
 82   6  25  14  4.8  US-10-098-263B-36091   Sequence 36091, A
 83   6  25  14  4.8  US-10-098-263B-45424   Sequence 45424, A
c84   6  25  14  4.8  US-10-098-263B-59601   Sequence 59601, A
c85   6  25  14  4.8  US-10-098-263B-59602   Sequence 59602, A
c86   6  25  14  4.8  US-10-098-263B-85954   Sequence 85954, A
c87   6  25  14  4.8  US-10-098-263B-86261   Sequence 86261, A
 88   6  25  14  4.8  US-10-098-263B-97035   Sequence 97035, A
 89   6  25  14  4.8  US-10-098-263B-97036   Sequence 97036, A
 90   6  25  14  4.8  US-10-098-263B-109737  Sequence 109737,
 91   6  25  14  4.8  US-10-098-263B-130586  Sequence 130586,
 92   6  28   9  4.8  US-09-751-561-55       Sequence 55, Appli
 93   6  28   9  4.8  US-09-852-000-2        Sequence 2, Appli
c94   6  28   9  4.8  US-09-852-000-3        Sequence 3, Appli
 95   6  28   9  4.8  US-09-852-000-6        Sequence 6, Appli
 96   6  28   9  4.8  US-09-852-000-7        Sequence 7, Appli
c97   6  28   9  4.8  US-09-852-000-8        Sequence 8, Appli
c98   6  28   9  4.8  US-09-852-000-9        Sequence 9, Appli
c99   6  30   8  4.8  US-08-805-813-6        Sequence 6, Appli
c100  6  30  14  4.8  US-10-085-906-97       Sequence 97, Appl
```

ALIGNMENTS

RESULT 1
US-08-860-844-18
; Sequence 18, Application US/08860844
; Publication No. US2003C104361A1
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF NUCLEIC ACIDS WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Saliwanchik & Saliwanchik
;   STREET: 2421 N.W. 41st St., Suite A-1
;   CITY: Gainesville
;   STATE: Florida
;   COUNTRY: USA
;   ZIP: 32606
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/860,844
;   FILING DATE: 09-JUN-1997
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/353,476
;   FILING DATE: 09-DEC-1994
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Saliwanchik, David R.
;   REGISTRATION NUMBER: 35,746
;   REFERENCE/DOCKET NUMBER: GP-100C1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (352) 375-8100
;   TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 33 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
US-08-860-844-18

```
Alignment Scores:
Pred. No.:              47                  Length:        33
Score:                  7.00               Matches:        7
Percent Similarity:     100.00%            Conservative:   0
Best Local Similarity:  100.00%            Mismatches:     0
Query Match:            5.65%              Indels:         0
DB:                     8                  Gaps:           0
```

US-09-484-577A-4 (1-124) x US-08-860-844-18 (1-33)

```
Qy   58 ValAlaTrpAlaGlyLeuGly 64
           |||||||||||||||||||
Db    2 GTGGCCTGGGCGGGACTGGGG 22
```

RESULT 2
US-10-407-543-18
; Sequence 18, Application US/10407543
; Publication No. US20031017S789A1
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
;   Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF NUCLEIC ACIDS WITH A
;   SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Saliwanchik & Saliwanchik
;   STREET: 2421 N.W. 41st St., Suite A-1
;   CITY: Gainesville
;   STATE: Florida
;   COUNTRY: USA
;   ZIP: 32606
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/407,543
;   FILING DATE: 03-Apr-2003
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/08/860,844
;   FILING DATE: 09-JUN-1997
;   APPLICATION NUMBER: 08/353,476
;   FILING DATE: 09-DEC-1994
; ATTORNEY/AGENT INFORMATION:
;   NAME: Saliwanchik, David R.
;   REGISTRATION NUMBER: 35,746
;   REFERENCE/DOCKET NUMBER: GP-100C1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (352) 375-8100
;   TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 33 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-407-543-18

Alignment Scores:
Pred. No.:   47   Length: 33
Score: 7.00   Matches: 7
Percent Similarity: 100.00%   Conservative: 0
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       October 23, 2003, 16:14:18 ; Search time 223 Seconds
                                  (without alignments) run
                                  6025.033 Million cell updates/sec

Title:        US-09-484-577A-3
Perfect score: 501
Sequence:     1 acctccagcctctcaccga........cagaatcggcgatgatcat 501

Scoring table: OLIGO_NUC
              Gapop 60.0 , Gapext 60.0

Searched:     1792395 seqs, 1340900451 residues

Word size :   0

Total number of hits satisfying chosen parameters:     1468544

Minimum DB seq length: 0
Maximum DB seq length: 250

Post-processing: Listing first 100 summaries

Database :    Published Applications NA:*
              1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
              2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
              3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
              4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
              5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
              6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
              7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
              8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
              9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
             10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
             11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
             12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
             13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
             14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
             15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
             16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
             17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | % Score | Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4.2 | 21 | 24 | 10 | US-09-928-457-55 | Sequence 55, Appl |
| 2 | 4.2 | 21 | 24 | 14 | US-10-193-451A-13 | Sequence 13, Appl |
| 3 | 4.2 | 21 | 28 | 9 | US-09-751-561-55 | Sequence 55, Appl |
| 4 | 4.0 | 20 | 23 | 9 | US-09-755-398A-2 | Sequence 2, Appli |
| 5 | 4.0 | 20 | 24 | 9 | US-09-751-561-1 | Sequence 15, Appli |
| 6 | 4.0 | 20 | 24 | 9 | US-09-751-561-15 | Sequence 15, Appl |
| 7 | 4.0 | 20 | 24 | 10 | US-09-928-457-58 | Sequence 58, Appl |
| 8 | 4.0 | 20 | 24 | 11 | US-09-989-364-1 | Sequence 1, Appli |
| 9 | 4.0 | 20 | 24 | 11 | US-09-989-364-5 | Sequence 5, Appl |
| 10 | 4.0 | 20 | 24 | 11 | US-09-989-364-19 | Sequence 19, Appl |
| 11 | 4.0 | 20 | 24 | 14 | US-10-193-451A-25 | Sequence 25, Appl |
| 12 | 4.0 | 20 | 24 | 14 | US-10-001-670-27 | Sequence 27, Appl |
| 13 | 4.0 | 20 | 24 | 14 | US-10-001-670-29 | Sequence 29, Appl |
| 14 | 4.0 | 20 | 24 | 14 | US-10-001-670-31 | Sequence 31, Appl |
| 15 | 4.0 | 20 | 24 | 14 | US-10-001-670-33 | Sequence 33, Appl |
| 16 | 4.0 | 20 | 24 | 14 | US-10-001-670-53 | Sequence 53, Appl |

| No. | Length | Score | ID | DB | Sequence |
|---|---|---|---|---|---|
| 17 | 20 | 4.0 | 24 | 14 | US-10-001-670-67 | Sequence 67, Appl |
| 18 | 20 | 4.0 | 24 | 14 | US-10-001-670-107 | Sequence 107, App |
| 19 | 20 | 4.0 | 24 | 14 | US-10-001-670-109 | Sequence 109, App |
| 20 | 20 | 4.0 | 48 | 9 | US-09-751-561-57 | Sequence 57, Appl |
| 21 | 20 | 4.0 | 53 | 9 | US-09-751-561-59 | Sequence 59, Appl |
| 22 | 19 | 3.8 | 23 | 14 | US-10-191-438-2 | Sequence 2, Appli |
| 23 | 19 | 3.8 | 24 | 14 | US-10-191-438-4 | Sequence 4, Appli |
| 24 | 19 | 3.8 | 24 | 8 | US-08-424-550B-1 | Sequence 1, Appli |
| 25 | 19 | 3.8 | 24 | 9 | US-09-751-797-1 | Sequence 1, Appli |
| 26 | 19 | 3.8 | 24 | 9 | US-09-930-334-7 | Sequence 7, Appl |
| 27 | 19 | 3.8 | 24 | 12 | US-10-096-534-69 | Sequence 69, Appl |
| 28 | 19 | 3.8 | 24 | 12 | US-10-134-345-7 | Sequence 7, Appli |
| 29 | 19 | 3.8 | 24 | 12 | US-10-348-190-22 | Sequence 22, Appl |
| 30 | 19 | 3.8 | 24 | 12 | US-10-160-237-11 | Sequence 11, Appl |
| 31 | 19 | 3.8 | 24 | 13 | US-10-032-626-7 | Sequence 7, Appli |
| 32 | 19 | 3.8 | 24 | 13 | US-10-067-813-6 | Sequence 6, Appl |
| 33 | 19 | 3.8 | 24 | 13 | US-10-085-108-11 | Sequence 11, Appl |
| 34 | 19 | 3.8 | 24 | 13 | US-10-235-264-4 | Sequence 4, Appli |
| 35 | 19 | 3.8 | 24 | 14 | US-10-193-451A-7 | Sequence 7, Appli |
| 36 | 19 | 3.8 | 24 | 14 | US-10-193-451A-19 | Sequence 19, Appl |
| 37 | 19 | 3.8 | 60 | 12 | US-09-908-975-5483 | Sequence 5483, Ap |
| 38 | 19 | 3.8 | 60 | 12 | US-09-908-975-18740 | Sequence 18740, A |
| 39 | 19 | 3.8 | 124 | 14 | US-10-191-438-3 | Sequence 3, Appli |  c
| 40 | 17 | 3.4 | 180 | 9 | US-09-791-244-3 | Sequence 3, Appli |
| 41 | 17 | 3.4 | 20 | 9 | US-09-758-735-8 | Sequence 8, Appli |
| 42 | 17 | 3.4 | 20 | 9 | US-09-758-735-12 | Sequence 12, Appl |  c
| 43 | 17 | 3.4 | 60 | 12 | US-09-908-975-18749 | Sequence 18749, A |
| 44 | 17 | 3.4 | 103 | 10 | US-09-954-456-300 | Sequence 300, App |  c
| 45 | 17 | 3.4 | 103 | 10 | US-09-954-456-1201 | Sequence 1201, Ap |
| 46 | 17 | 3.4 | 103 | 10 | US-09-880-107-486 | Sequence 486, App |
| 47 | 17 | 3.4 | 103 | 10 | US-09-954-531-1007 | Sequence 1007, Ap |
| 48 | 17 | 3.4 | 103 | 10 | US-09-873-319-94 | Sequence 94, Appl |
| 49 | 17 | 3.4 | 103 | 12 | US-09-960-706-145 | Sequence 145, App |
| 50 | 16 | 3.2 | 17 | 12 | US-10-230-006-57 | Sequence 57, Appl |
| 51 | 16 | 3.2 | 17 | 12 | US-10-230-006-2103 | Sequence 2103, Ap |  c
| 52 | 15 | 3.0 | 17 | 12 | US-10-230-006-568 | Sequence 568, App |  c
| 53 | 15 | 3.0 | 17 | 12 | US-10-230-006-1272 | Sequence 1272, Ap |  c
| 54 | 15 | 3.0 | 19 | 12 | US-10-226-992-23 | Sequence 23, Appl |
| 55 | 15 | 3.0 | 19 | 12 | US-10-226-992-106 | Sequence 106, App |  c
| 56 | 15 | 3.0 | 25 | 14 | US-10-098-263B-39494 | Sequence 39494, A |
| 57 | 15 | 3.0 | 25 | 14 | US-10-098-263B-96190 | Sequence 96190, A |
| 58 | 15 | 3.0 | 87 | 9 | US-09-864-761-29443 | Sequence 29443, A |  c
| 59 | 15 | 3.0 | 179 | 9 | US-09-864-761-27258 | Sequence 27258, A |
| 60 | 15 | 3.0 | 237 | 10 | US-09-728-444-257 | Sequence 257, App |  c
| 61 | 15 | 3.0 | 237 | 10 | US-09-796-692-3378 | Sequence 3378, Ap |  c
| 62 | 15 | 3.0 | 237 | 14 | US-10-040-862-3378 | Sequence 3378, Ap |  c
| 63 | 15 | 3.0 | 248 | 10 | US-09-878-965-3631 | Sequence 3631, Ap |
| 64 | 15 | 3.0 | 15 | 10 | US-09-878-574-11290 | Sequence 11290, A |
| 65 | 14 | 2.8 | 14 | 9 | US-09-755-398A-4 | Sequence 4, Appli |
| 66 | 14 | 2.8 | 25 | 14 | US-10-098-263B-75929 | Sequence 75929, A |  c
| 67 | 14 | 2.8 | 14 | 14 | US-10-098-263B-106023 | Sequence 106023, A |  c
| 68 | 14 | 2.8 | 29 | 12 | US-10-301-516-6 | Sequence 6, Appli |
| 69 | 14 | 2.8 | 33 | 9 | US-09-861-097-8 | Sequence 8, Appli |
| 70 | 14 | 2.8 | 33 | 11 | US-09-861-098-8 | Sequence 8, Appli |
| 71 | 14 | 2.8 | 33 | 12 | US-09-861-012-8 | Sequence 8, Appli |
| 72 | 14 | 2.8 | 34 | 11 | US-10-051-989-8 | Sequence 67, Appl |
| 73 | 14 | 2.8 | 38 | 11 | US-09-988-115A-67 | Sequence 141, App |
| 74 | 14 | 2.8 | 38 | 12 | US-10-280-137-141 | Sequence 141, App |  c
| 75 | 14 | 2.8 | 38 | 12 | US-10-280-137-142 | Sequence 142, App |  c
| 76 | 14 | 2.8 | 40 | 12 | US-10-280-137-147 | Sequence 147, App |
| 77 | 14 | 2.8 | 40 | 12 | US-10-280-137-148 | Sequence 148, App |
| 78 | 14 | 2.8 | 60 | 12 | US-10-202-896-47 | Sequence 47, Appl |
| 79 | 14 | 2.8 | 85 | 10 | US-10-202-896-53 | Sequence 53, Appl |
| 80 | 14 | 2.8 | 85 | 10 | US-09-975-132A-8 | Sequence 8, Appli |
| 81 | 14 | 2.8 | 88 | 10 | US-09-864-761-20883 | Sequence 20883, A |  c
| 82 | 14 | 2.8 | 88 | 10 | US-09-796-692-5003 | Sequence 5003, Ap |  c
| 83 | 14 | 2.8 | 105 | 14 | US-10-040-862-5003 | Sequence 5003, Ap |  c
| 84 | 14 | 2.8 | 119 | 10 | US-09-998-598-1860 | Sequence 1860, Ap |
| 85 | 14 | 2.8 | 137 | 9 | US-09-864-761-22266 | Sequence 22266, A |  c
| 86 | 14 | 2.8 | 139 | 9 | US-09-864-761-21451 | Sequence 21451, A |  c
| 87 | 14 | 2.8 | 143 | 9 | US-09-864-761-21357 | Sequence 21357, A |  c
| 88 | 14 | 2.8 | 143 | 9 | US-09-923-876-1171 | Sequence 1171, Ap |  c
| 89 | 14 | 2.8 | 143 | 9 | US-09-864-761-23978 | Sequence 23978, A |  c

```
c   90   14   2.8   156    9   US-09-736-969A-65      Sequence 65, Appl
c   91   14   2.8   157   10   US-09-878-574-8329     Sequence 8329, Ap
c   92   14   2.8   179   10   US-09-867-701-9404     Sequence 9404, Ap
    93   14   2.8   180   12   US-10-029-386-19181    Sequence 19181, A
    94   14   2.8   181    9   US-09-864-761-21030    Sequence 21030, A
c   95   14   2.8   190   10   US-09-960-352-9546     Sequence 9546, Ap
c   96   14   2.8   198   10   US-09-796-692-7865     Sequence 7865, Ap
c   97   14   2.8   198   10   US-09-796-692-8040     Sequence 8040, Ap
c   98   14   2.8   198   14   US-10-040-862-7865     Sequence 7865, Ap
c   99   14   2.8   198   14   US-10-040-862-8040     Sequence 8040, Ap
   100   14   2.8   204   10   US-09-738-626-1469     Sequence 1469, Ap
```

ALIGNMENTS

RESULT 1
US-09-928-457-55
; Sequence 55, Application US/03928457
; Patent No. US20020164603A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA, specific proteins and peptides
; TITLE OF INVENTION: of the Neisseria meningitidis species bacteria, method
; TITLE OF INVENTION: for obtaining them and their biological application.
; NUMBER OF SEQUENCES: 99
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (OEB)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/928,457
; FILING DATE: 2001-08-14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/214,759
; FILING DATE: 199-12-10
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-928-457-55

Query Match   4.2%; Score 21; DB 1C; Length 24;
Best Local Similarity   100.0%; Pred. No. 0.23;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1   ACTCTCCAGCCTCTCACCGAG 21
          |||||||||||||||||||||
Db    4   ACTCTCCAGCCTCTCACCGAG 24
```

RESULT 2
US-10-193-451A-13
; Sequence 13, Application US/10193451A
; Publication No. US2003009269A1
; GENERAL INFORMATION:
; APPLICANT: CULLIS, CHRISTOPHER A.
; APPLICANT: RADEMAN, SAMANTHA
; APPLICANT: KUNERT, KARL
; TITLE OF INVENTION: METHOD FOR FINDING GENETIC MARKERS
; TITLE OF INVENTION: OF SOMACLONAL VARIATION
; FILE REFERENCE: 02831560005
; CURRENT APPLICATION NUMBER: US/10/193,451A
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/292,646
; PRIOR FILING DATE: 1999-04-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: WordPerfect for Windows v. 7.0
; SEQ ID NO 13
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
US-10-193-451A-13

Query Match   4.2%; Score 21; DB 14; Length 24;
Best Local Similarity   100.0%; Pred. No. 0.23;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy    1   ACTCTCCAGCCTCTCACCGAG 21
          |||||||||||||||||||||
Db    4   ACTCTCCAGCCTCTCACCGAG 24
```

RESULT 3
US-09-751-561-55
; Sequence 55, Application US/09751561
; Patent No. US20010007985A1
; GENERAL INFORMATION:
; APPLICANT: Rothberg, Jonathan
; APPLICANT: Deem, Michael
; APPLICANT: Simpson, John
; TITLE OF INVENTION: Method for the Determination and
; TITLE OF INVENTION: Classification of DNA Sequences in a Sample Without
; TITLE OF INVENTION: Sequencing
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie and Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/751,561
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/547,214
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7934-015-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)-790-9090
; TELEFAX: (212)-869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-751-561-55

Query Match   4.2%; Score 21; DB 9; Length 28;
Best Local Similarity   100.0%; Pred. No. 0.22;
Matches   21; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
86   6  4.8  US-08-663-823B-60   2   Sequence 60, Appl
87   6  4.8  US-08-663-823B-62   2   Sequence 62, Appl
88   6  4.8  US-08-993-118-11    2   Sequence 11, Appl
89   6  4.8  US-08-748-130-23    2   Sequence 23, Appl
90   6  4.8  US-08-845-528C-11   3   Sequence 11, Appl
91   6  4.8  US-08-874-825-27    3   Sequence 27, Appl
92   6  4.8  US-08-874-825-29    3   Sequence 29, Appl
93   6  4.8  US-08-874-825-31    3   Sequence 31, Appl
94   6  4.8  US-08-874-825-33    3   Sequence 33, App-
95   6  4.8  US-08-874-825-53    3   Sequence 53, Appl
96   6  4.8  US-08-874-825-67    3   Sequence 67, Appl
97   6  4.8  US-08-874-825-107   3   Sequence 107, App
98   6  4.8  US-08-874-825-109   3   Sequence 109, App
99   6  4.8  US-08-663-824-27    3   Sequence 27, Appl
100  6  4.8  US-09-663-824-29    3   Sequence 29, Appl
```

ALIGNMENTS

RESULT 1
US-08-353-476-18
; Sequence 18, Application US/08353476
; Patent No. 5871902
; GENERAL INFORMATION:
; APPLICANT: Weininger, Susan
; APPLICANT: Weininger, Arthur M
; TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
; TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st St., Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,476
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: GP-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-353-476-18

```
Alignment Scores:
Pred. No.:              21.4        Length:         33
Score:                  7.00        Matches:         7
Percent Similarity:     100.00%     Conservative:    0
Best Local Similarity:  100.00%     Mismatches:      0
Query Match:            5.65%       Indels:          0
DB:                     2           Gaps:            0
```

US-09-484-577A-4 (1-124) x US-08-353-476-18 (1-33)

```
Qy  58 ValAlaTrpAlaGlyLeuGly 64
        |||||||||||||||||||||
Db   2 GTGGCCTGGGCGGGACTGGGG 22
```

RESULT 2
US-08-989-394-17/c
; Sequence 17, Application US/08989394
; Patent No. 5994136
; GENERAL INFORMATION:
; APPLICANT: Naldini, Luigi
; APPLICANT: Dull, Thomas
; APPLICANT: Farson, Deborah A.
; APPLICANT: Witt, Rochelle
; TITLE OF INVENTION: METHOD AND MEANS FOR PRODUCING HIGH
; TITLE OF INVENTION: TITER, SAFE, RECOMBINANT LENTIVIRUS VECTORS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,394
; FILING DATE: 12-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Dean H.
; REGISTRATION NUMBER: 33,981
; REFERENCE/DOCKET NUMBER: A7086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-989-394-17

```
Alignment Scores:
Pred. No.:              27          Length:         42
Score:                  7.00        Matches:         7
Percent Similarity:     100.00%     Conservative:    0
Best Local Similarity:  100.00%     Mismatches:      0
Query Match:            5.65%       Indels:          0
DB:                     2           Gaps:            0
```

US-09-484-577A-4 (1-124) x US-08-989-394-17 (1-42)

```
Qy  58 ValAlaTrpAlaGlyLeuGly 64
        |||||||||||||||||||||
Db  37 GTGGCCTGGGCGGGACTGGGG 17
```

RESULT 3
US-09-271-365-17/c
; Sequence 17, Application US/09271365
; Patent No. 6155782
; GENERAL INFORMATION:
; APPLICANT: Naldini, Luigi
; APPLICANT: Dull, Thomas
; APPLICANT: Farson, Deborah A.
; APPLICANT: Witt, Rochelle
; TITLE OF INVENTION: METHOD AND MEANS FOR PRODUCING HIGH
;   TITER, SAFE, RECOMBINANT LENTIVIRUS VECTORS